INVESTOR TYPES AND RISK AVERSION: A DATA SCIENCE APPROACH TO FINANCIAL DECISION-MAKING

A THESIS
SUBMITTED TO ABDULLAH GÜL UNIVERSITY
SOCIAL SCIENCES INSTITUTE
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE

By
Beyza Aytemur
September, 2025
Kayseri

Beyza Aytemur
INVESTOR TYPES AND RISK AVERSION: A DATA SCIENCE APPROACH TO FINANCIAL DECISION- MAKING
AGU
2025

INVESTOR TYPES AND RISK AVERSION: A DATA SCIENCE APPROACH TO FINANCIAL DECISION- MAKING

A THESIS

SUBMITTED TO ABDULLAH GÜL UNIVERSITY

SOCIAL SCIENCES INSTITUTE

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF SCIENCE

By

Beyza Aytemur

September, 2025

Kayseri

SCIENTIFIC ETHICS COMPLIANCE

I hereby declare that all information in this document has been obtained in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name-Surname: Beyza Aytemur

Signature :

REGULATORY COMPLIANCE

M.Sc. thesis titled “Investor Types and Risk Aversion: A Data Science Approach to Financial Decision- Making” has been prepared in accordance with the Graduate Thesis Preparation Guidelines of the Abdullah Gül University, Social Sciences Institute.

Prepared By	Advisor
Beyza Aytemur	Prof. Dr. Erk Hacıhasanoğlu

Head of the Data Science for Business and Economics Program

Assoc. Prof. Dr. Umut Türk

ACCEPTANCE AND APPROVAL

M.Sc. thesis titled "Investor Types and Risk Aversion: A Data Science Approach to Financial Decision- Making" and prepared by Beyza Aytemur has been accepted by the jury in the Data Science Graduate Program Graduate Program at Abdullah Gül University, Social Sciences Institute.

……….. /……….. / ………..

(Thesis Defense Exam Date)

JURY:

Advisor : Prof. Dr. Erk Hacihasanoğlu ....…………………

Member : Doç. Dr. Adil Oran.…………………..................................

Member : Prof. Dr. Onur Gözbaşı………………..............................

APPROVAL:

The acceptance of this M.Sc. thesis has been approved by the decision of the Abdullah Gül University, Social Sciences Institute, Management Board dated ….. /….. / ……….. and numbered ……………..……. .

……….. /……….. /
………..

Assoc. Prof. Dr.Umut Türk

Director of Social Sciences Institute

Name Surname : Beyza Aytemur

Program : Data Science (Master of Science)

Advisor : Prof. Dr. Erk Hacıhasanoğlu

Thesis Title : Investor Types and Risk Aversion: A Data Science Approach to Financial Decision- Making

Date and Place : September, 2025 – Kayseri, Turkey

ABSTRACT

This study examines the manifestation of behavioral biases in investor decision-making through a quantitative data science approach using the FAR-Trans dataset. We developed empirical measures for three key behavioral finance constructs (loss aversion, herding behavior, and overconfidence) derived from established theoretical frameworks and applied them to classify investors into distinct behavioral groups. A decision tree-based multi-class classification model was employed to predict behavioral bias categories, demonstrating exceptional predictive performance with 96% accuracy and macro-average ROC-AUC scores of 0.99 validated through cross-validation analysis. Statistical significance tests and visual analyses confirm that each behavioral group exhibits clearly distinguishable patterns in their respective dominant metrics, with loss-averse investors displaying higher tendencies to maintain losing positions, herding investors showing increased propensity to follow popular assets, and overconfident investors engaging in more frequent trading activities. These findings provide robust empirical evidence for the effective identification and classification of behavioral biases from transactional financial data and demonstrate the practical utility of machine learning techniques for understanding psychological factors in financial markets.

Keywords: Behavioral finance, Investor types, Data science, Machine learning, Decision tree algorithm.

Ad Soyad : Beyza Aytemur

Anabilim Dalı, Program : Veri Bilimi (Yüksek Lisans)

Tez Danışmanı : Prof. Dr. Erk Hacıhasanoğlu

Tez Başlığı : Yatırımcı Tipleri ve Riskten Kaçınma: Finansal Karar Vermede Veri Bilimi Yaklaşımı

Tarih ve Yer : Eylül, 2025 – Kayseri, Türkiye

ÖZET

Bu çalışma, yatırımcıların karar alma süreçlerindeki davranışsal önyargıların ortaya çıkışını, FAR-Trans veri seti kullanılarak nicel bir veri bilimi yaklaşımıyla incelemektedir. Davranışsal finansın üç temel kavramı olan kayıptan kaçınma, sürü davranışı ve aşırı özgüven için kuramsal çerçevelere dayanan ampirik ölçütler geliştirilmiş ve bunlar yatırımcıların farklı davranış gruplarına sınıflandırılmasında kullanılmıştır. Davranışsal önyargı kategorilerini tahmin etmek amacıyla karar ağacı tabanlı çok sınıflı bir sınıflandırma modeli uygulanmış ve bu model çapraz doğrulama analizleriyle %96 doğruluk ve 0.99 makro-ortalama ROC-AUC skoru ile oldukça yüksek bir öngörü performansı sergilemiştir. İstatistiksel anlamlılık testleri ve görsel analizler, her davranışsal grubun kendi baskın metriğinde net bir şekilde ayrıştığını doğrulamaktadır; kayıptan kaçınan yatırımcılar kayıplı pozisyonları elde tutma eğilimini daha yüksek gösterirken, sürü davranışı sergileyen yatırımcılar popüler varlıkları takip etme eğiliminde, aşırı özgüvenli yatırımcılar ise daha sık işlem yapma eğilimindedir. Bu bulgular, finansal işlem verilerinden davranışsal önyargıların etkin şekilde tanımlanıp sınıflandırılabileceğine dair güçlü ampirik kanıt sunmakta ve makine öğrenmesi tekniklerinin finansal piyasalarda psikolojik faktörleri anlamada pratik faydasını ortaya koymaktadır.

Anahtar kelimeler: Davranışsal finans, Yatırımcı tipleri, Veri bilimi, Makine öğrenmesi, Karar ağacı algoritması.

ACKNOWLEDGEMENTS

This thesis was prepared as part of my master's degree at Abdullah Gül University, Faculty of Social Sciences, Department of Data Science. I would like to thank my esteemed advisor Prof. Dr. Erk Hacıhasanoğlu, who guided me with his valuable knowledge and experience at every stage of the study and never withheld his support.

I am also grateful to my father, my friends Meltem, Maham, Kutalmış, Mustafa and Sergen, whose moral support I always felt during the master's degree process.

I am grateful to everyone who contributed to this process.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

EMT	Efficient Market Theory
EUT	Expected Utility Theory
PT	Prospect Theory

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Financial decision making has been a subject of debate from past to present. It has been a matter of curiosity what people pay attention to when making financial decisions and which factors affect them. Different hypotheses and theories have been developed on this subject.

Traditional finance typically models investors as rational utility maximizers, often summarized as the *homo economicus* assumption. In this view, individuals are fully informed, can process information without bias, and act consistently to maximize utility or profit, a conception that can be traced to early discussions of the “economic man” (Mill, 1836) and that was later formalized in mathematical terms (Edgeworth, 1881). Under this framework, market participants are assumed to analyze available information comprehensively before making investment choices and to behave in accordance with expected utility maximization. Early definitions also emphasized economics as choice under scarcity and constraints (Robbins, 1932), and the maximizing framework was extended to a wide range of human behavior (Becker, 1976). In contrast, the bounded rationality perspective argues that real decision makers face cognitive and informational limits, so optimization is often replaced by satisficing rules and heuristics (Simon, 1955).

Efficient Market Theory (EMT) builds on the rationality assumption by positing informational efficiency—namely, that asset prices fully and promptly incorporate all available information. Although EMT is commonly associated with Fama (1970), earlier arguments anticipated its logic: Bachelier (1900) modeled speculative prices as a fair game with random changes, Kendall (1953) documented the near-random behavior of price series, and Samuelson (1965) showed that properly anticipated prices follow a martingale. In this view, investors cannot consistently earn returns above the market without assuming additional risk, since new information is rapidly embedded in prices (Fama, 1970; Malkiel, 1973). The implication is that, under rational expectations and competitive markets, price anomalies should not persist because information is quickly reflected in asset values—albeit with the caveat that perfectly costless, perfectly efficient markets are infeasible (Grossman & Stiglitz, 1980).

However, the limitations of traditional finance theories have become increasingly evident, especially in light of anomalies and psychological factors that contradict the assumption of rationality. For instance, behavioral finance scholars argue that investors often behave irrationally due to cognitive biases, emotions, and social influences, which can lead to systematic errors in the decision-making process (Almansour and Arabyat, 2017; Kobiyh et al., 2023). Such inconsistencies have led to a reconsideration of the rigid assumptions of traditional finance, especially during periods of market stress when rational human understanding cannot accurately predict investor behavior.

Other criticisms of traditional finance stem from the observation that many investment decisions are influenced by heuristics and biases that can deviate significantly from rational decision-making. For example, tendencies and biases such as overconfidence, loss aversion, herd psychology, etc. can lead to behaviors that contradict the predictions of EMT, resulting in market mispricing and inefficiencies (Cheng, 2022). These factors indicate that there is a growing recognition that human psychology plays an important role in financial decision-making, challenging the idealized rational investor model.

In this context, the understanding that human beings cannot be fully rational and that their emotions, thoughts, experiences, psychology and environment are influential in decision-making has been adopted since the half of the 20th century. The studies that humans are far from being rational started with Herbert Simon's criticisms of rationality. In this context, Daniel Kahneman and Amos Tversky's “Prospect Theory: Decision Making Under Uncertainty” published in ‘Econometrica’ in 1979, laid the foundations of behavioral finance.

Behavioral finance has emerged as an important counterpoint to traditional finance theories, primarily due to criticisms of the assumptions of rationality and market efficiency underlying classical finance models. Traditional finance generally assumes that market participants are rational and make decisions based solely on available information to maximize their utility. EMT argues that security prices always reflect all available information and neglect psychological factors that may influence decision-making processes (Kumar & Goyal, 2016).

On the other hand, behavioral finance integrates concepts from psychology into the study of financial markets, recognizing that human behavior often deviates from the rational model (Yang, 2016). This field emphasizes that investors do not always act in their best financial interests and that cognitive biases, emotional reactions, and social factors significantly influence their decisions (Jaiyeoba et al., 2019). For example, situational conditions such as panic or euphoria can cause herding behavior, where investors collectively move in a certain direction, leading to market anomalies (Zhang and Zheng, 2015).

Moreover, behavioral finance challenges the rational actor model by promoting concepts such as “bounded rationality” and psychological biases that highlight the limits of decision-making among investors (Jaiyeoba et al., 2019). It examines how various biases such as overconfidence, loss aversion, agency, etc. affect investor behavior (Tekin, 2016). Research has shown that both institutional and individual investors often exhibit irrational behavior that traditional finance cannot adequately explain (Piras, 2012). Investors' bias to succumb to biases can lead to mispricing of assets, scenarios that classical finance theories cannot effectively predict (Yi et al., 2021).

Contemporary criticisms of traditional finance are supported by empirical studies that challenge the assumptions of market efficiency and rational behavior. Some scholars argue that traditional models fail to adequately explain market events such as bubbles or crashes, as observed during major financial crises (Cheng, 2022). This inadequacy has led to a reassessment of the fundamental principles of finance and the advocacy of methodologies that incorporate human psychology (Karsh, 2018). The inclusion of behavioral factors provides a more nuanced understanding of market dynamics and investor behavior by shifting the focus from individual rationality to social influences and networks (Zhang and Zheng, 2015).

Behavioral finance recognizes that human beings cannot have rational and complete information when making financial decisions, and even if they have information, their psychological state, tendencies and prejudices, social environment, etc. will have an impact on their decisions. In the financial field, disciplines such as psychology, sociology, anthropology, etc. are also included. Behavioral finance frequently uses statistical and quantitative analysis methods to examine the impact of individuals' irrational behavior and psychological biases on financial decision-making processes. Developments in data science, machine learning and artificial intelligence have facilitated the analysis of metadata. Moreover, these developments make it possible to develop more precise models by analyzing behavioral biases over large data sets. In particular, machine learning methods are effective in explaining individuals' financial behaviors such as investment preferences, risk perception and spending habits through clustering, classification and prediction.

Data science applications in financial decision-making processes have become widespread in areas such as customer segmentation, fraud detection and portfolio optimization. For example, clustering algorithms can be used to determine the investment profiles of individuals. In this way, financial institutions can offer more personalized products to their customers. In addition, techniques such as sentiment analysis can be used to predict market movements by analyzing social media and news data. These approaches combine the theoretical framework of behavioral finance with practical applications to develop more effective strategies.

K-Means and DBSCAN are used in this study as complementary and exploratory tools rather than the primary inference engine. K-Means partitions observations into k clusters by minimizing within-cluster variance; it is fast and scalable but requires choosing k and tends to favor roughly convex, similarly sized groups. DBSCAN is density-based, can discover arbitrarily shaped clusters, and labels sparse points as noise; it does not require a prespecified number of clusters but is sensitive to the choices of ε and min_samples. In our setting, both methods are applied to probe latent structure and to support interpretation, while all quantitative conclusions come from supervised models.

Prior research shows that unsupervised and supervised machine learning both contribute to financial analytics in tasks such as customer segmentation, anomaly detection, and risk grouping. In behavioral finance, clustering has been used to form interpretable investor groups for comparing outcomes, and theory clarifies how systematic biases shape decisions, which guides the design of behavior-aligned features. These strands motivate the use of clustering as an exploratory lens, while validated and decision-relevant predictions are produced by supervised classification.

Using the FAR-Trans dataset, we build a supervised pipeline that cleans transactions, engineers features grounded in the literature (for example trading activity and scale, loss-aversion proxies, and herding indicators), and trains cross-validated classifiers evaluated with accuracy, ROC-AUC, and confusion matrices. We conduct feature-importance analysis to identify the most informative signals. K-Means and DBSCAN are applied only to segment investors in an unsupervised manner and to run robustness checks on the supervised findings, not to draw statistical conclusions. By improving decision quality and enabling more productive, risk-aware participation in capital markets, the approach supports SDG 8 (Decent Work and Economic Growth). By packaging the workflow as a modular analytics layer that can be embedded into brokerage infrastructure, it advances SDG 9 (Industry, Innovation and Infrastructure).

2. LITERATURE REVIEW

2.1. Behavioral Finance Concept

Traditional finance posits expected-utility-maximizing investors who choose portfolios by trading off risk and return (von Neumann & Morgenstern, 1944). Variation in risk tolerance is not contrary to rational behavior, because differences in preferences and beliefs imply different optimal choices. Behavioral finance complements this view by showing how internal and external influences shape those inputs to choice, while decisions remain rational given perceived states and constraints (Shleifer, 2000).

In contrast to traditional finance theories, behavioral finance is an interdisciplinary field that examines the irrational decisions of individuals and their impact on markets. Traditional finance models assume that people always act in a rational and self-interest maximizing manner (Fama, 1970). However, behavioral finance investigates how psychological factors (e.g., biases, emotions, cognitive constraints) influence investment decisions (Kahneman and Tversky, 1979).

Behavioral finance is an interdisciplinary field that combines two different disciplines, finance and psychology, in an effort to explain why and how people can make irrational decisions when it comes to saving, spending and investing money (Belsky and Gilovich, 2000). Therefore, why people do not make rational financial decisions is also included in this field. According to Shefrin (2002, p. 3), behavioral finance is an approach that examines how psychological factors affect financial decision-making processes and market outcomes.

However, in the field of behavioral finance, social sciences such as sociology, anthropology, etc. are used in addition to psychology. In traditional economics and finance, the decisions made by individuals are examined with the findings obtained from these disciplines and try to explain how people's psychology, thoughts and behaviors affect their financial decisions. According to Baltussen (2009), behavioral finance examines the behavior of financial markets based on the sciences that study human behavior. By identifying how emotional and cognitive errors affect investors and their decision-making processes, behavioral finance attempts to explain irrational investor behavior that is not found in traditional rational models.

Behavioral finance can be defined as an attitude that involves planning about the use of money, recording the use of money, revaluing fixed costs, and recording all financial transactions (Xiao, 2016).

According to Hirshleifer (2015, p. 147), behavioral finance systematically models individuals' decisions shaped by limited rationality, emotional reactions and social influences. In this context, an individual's financial behavior is shaped by a combination of innate tendencies, knowledge and skills acquired from family, environment and experiences, and psychological factors (Dew and Xiao, 2011).

Behavioral finance is a discipline that examines human tendencies and errors in making financial decisions and the functioning of financial markets. This field analyzes factors such as the risk-taking tendencies of individuals and investors, their incentives to avoid losses, the way they access information and the effects of the social environment (Cialdini, 1998).

It is possible to multiply the definitions of behavioral finance in the literature. As a matter of fact, Pailer and Yıldız (2021), who examined the definitions of the concept of behavioral finance, expressed the differences among these definitions as follows:

- Evaluation of mainstream theories of thought in finance together with psychology
- It is the branch of science that tries to explain the causes of anomalies in financial markets.
- The branch of science that tries to obtain a significant result by investigating the psychological errors made by investors in the decision-making problem.
- The branch of science that studies the irrational decisions of normal people
- It is an alternative approach to the Efficient Market Theory.
- It is the study of situations in which decision-makers consider satisfaction rather than maximum utility.

Behavioral finance rejects the 'rational human' assumption of standard finance theory and draws on psychology, sociology and neuroscience to analyze the systematic errors of investors (Statman, 1999, p. 19). Behavioral finance is a field that aims to combine these disciplines with traditional economics and finance to explain why people make irrational financial decisions. Unlike traditional finance, behavioral finance assumes that there are limits to arbitrage and that not all people are rational (Ricciardi and Simon, 2000). It is also a discipline that examines how people actually behave in financial markets. This behavior often deviates from what rational models predict (Shiller, 2000, p. 83).

Behavioral finance argues that some financial problems can be solved more easily by using a model of human behavior that is not fully rational rather than a model of rational human behavior. If people are not rational as predicted in traditional financial models, what happens when the assumption of rationality is removed or treated more flexibly is the subject of behavioral finance (Barberis and Thaler, 2003).

At the heart of behavioral finance is the idea that psychological factors can make financial analyses produce findings of values closer to reality and that the findings obtained will improve institutional understanding in the field of finance and suggest better policies by obtaining more accurate predictions closer to reality (Camerer and Loewenstein, 2004).

Thaler (2015) explains that the main goal of behavioral finance is to understand why people systematically deviate from rational models. In addition, behavioral finance offers various assumptions and models to explain market anomalies, understand investor psychology and detect systematic errors in financial decision-making processes (Barberis & Thaler, 2003).

One of the main subject areas of behavioral finance is the study of human behavior within a social science framework. In this case, taking psychological factors into account while investigating human behavior is an important factor in analyzing people's decision-making processes. Behavioral finance has dealt with the behavior of the individual within the framework of rationality and has included the factors affecting the psychology of the individual in economic analysis. The inclusion of factors affecting the psychology of the individual in economic analysis leads to more realistic attitudes and results.

2.2. Emergence and Development of The Concept of Behavioral Finance

Behavioral finance is a relatively new area that emerged in the mid-20th century. This understanding assumes that people cannot always be rational when making economic decisions and that these decisions will be influenced by their psychology, emotions, experiences, environment and judgments. In economic models based on the traditional understanding of economics, the basic assumption is that people are rational in their economic and financial affairs and that they have all the relevant information. This understanding was also dominant in the early 20th century.

However, until the mid-19th century, when the classical economic approach was dominant, the idea that individuals' psychological states affect their financial behavior was included. In Adam Smith's book “The Wealth of Nations”, he states the psychological foundations behind human behavior. In particular, he deals with concepts such as pride, shame, insecurity and selfishness.

In the neo-classical period that emerged in later times, the sciences of economics and psychology moved away from each other and focused on the concept of homo economicus. According to this understanding, the individual who calculates, who is emotionless when it comes to economic decisions, and who only tries to maximize utility appears as “Homo Economicus” (Thaler and Sendhill, 2000).

“Homo Economicus” is an individual who exhibits rational behavior, is self-interested (trying to maximize his/her own interest), knows what he/she wants (stable in his/her preferences) and has full information in all situations (Scheneider, 2010). According to this understanding, human beings are considered rational.

In his “General Theory” (1936), John Maynard Keynes emphasized that financial decisions are not based solely on rational calculations. With the concept of “animal motives”, he argued that investor behavior is influenced by psychological and emotional factors (Keynes, 1936/2018, p. 161).

George Katona, working with researchers at the University of Michigan, pioneered an empirically grounded behavioral approach. In Psychological Analysis of Economic Behavior he argued that economic processes depend directly on human behavior and that psychological concepts should be incorporated into economic analysis, drawing on Gestalt psychology’s idea of perceptual wholeness, where the whole is not merely the sum of its parts and cognition shapes how individuals perceive situations (Katona, 1951). As an early policy application, during World War II he analyzed inflation and argued that assessing the ongoing inflation process requires measuring households’ inflation expectations (Katona, 1942). He also showed that learning and the surrounding environment shape these expectations, and he set out how psychological evidence can contribute to economic analysis (Katona, 1947, pp. 457–458).

The studies that human beings are far from being rational started with Herbert Simon's criticisms of rationality. In his work “Managerial Behavior” published in 1947, Herbert Simon, in contrast to the concept of the individual who maximizes the Utility of economic theories that traditional economics examined by returning to the basis of economic theories, drew attention to the variables that are effective in decision-making processes and put forward an idea in this course. He analyzed the behavior of individuals in society and placed the decision-making stage of the individual at the root of this research. At the same time, he argues that the physical conditions in which the individual lives and the character traits of the individual have an impact on their behavior. Simon, who says that the individual deviates from rationality in some situations, believes that when organizational goals are at the forefront, when the individual joins the group, he acts with the goals and wishes of the group he is a part of, not his own goals and wishes.

In his 1955 article “A Behavioral Model of Rational Choice”, Simon argued that the assumption of “economic man” in traditional economics needed a fundamental refinement. According to him, universal rationality is not possible since the individual's own experience and ability will prevent him from acting rationally. A large part of human behavior is determined by the emotions in which he or she finds himself or herself. Humans are rational to a certain extent and emotions affect the behavior of the individual. Since the individual's brain has limited computational capability, there has been a lack of information and uncertainty in the preference and decision process (Simon, 1955, 112-113).

In this context, Daniel Kahneman and Amos Tversky, building on Herbert Simon’s ideas about bounded rationality, introduced Prospect Theory in their 1979 Econometrica article “Prospect Theory: An Analysis of Decision under Risk.” The paper formalized reference dependence, loss aversion, and probability weighting, providing a psychological foundation for observed departures from expected utility theory and helping to establish behavioral finance as a research program. In 2002 Daniel Kahneman received the Nobel Prize in Economic Sciences for integrating insights from psychology into economics; Amos Tversky, who passed away in 1996, was not eligible for the prize. Following this work, behaviorally informed research in economics and finance expanded rapidly.

2.3. From Traditional Finance Theories to Behavioral Finance Theories

From past to present, many hypotheses and theories have been put forward on how individuals make decisions in financial markets. Under this heading, the most prominent of these theories will be analyzed.

2.3.1. Efficient Market Theory

The foundations of Efficient Market Theory (EMT) were laid by Eugene Fama’s (1970) “Efficient Capital Markets: A Review of Theory and Empirical Work.” In an informationally efficient market, prices fully reflect all available information. EMT therefore holds that, because information is already embedded in security values, the prices at which stocks and bonds trade today represent their fair value (Ricciardi and

Simon, 2000, pp. 1–2). As a result, investors cannot consistently outperform the market without assuming additional risk.

The validity of the EMT depends on many assumptions and these assumptions are related to the functioning of the market and the behavior of investors. In general, these assumptions are that the main objective of the investor is to maximize the utility of ultimate wealth (Barone, 2003). Investors make choices based on risk and return. Investors' risk and return expectations are homogeneous. Investors have the same information at the same time, that is, they think in a similar way (Taner and Kayalıdere, 2002, p. 2).

According to the EMT, it is assumed that investors have access to information at the same time and therefore no investor can earn abnormal returns. In addition, it is also assumed that security prices in efficient markets do not depend on past price movements, in other words, they are formed randomly (Malcıoğlu and Aydın, 2013). According to the theory, the true value and expected return of an asset should be calculated accurately using all available information (past earnings, financial statements, announcements of mergers, acquisitions, etc.) (Fama, 1976, p. 145).

2.3.2. Expected Utility Theory

One of the fundamental and most important theories of traditional finance is the Expected Utility Theory (EUT). It was first articulated by Bernoulli in 1738. This theory was developed and mathematically explained by two researchers, Jon Von Neumann and Oscar Morngenstren, in the “Theory of Games and Economic Behavior” published in 1944 (Sefil Tanrısever and Çilingiroğlu, 2011, pp. 249-250).

Bernoulli introduces the concept of diminishing marginal utility and its impact on decision making. Decisions were assumed to be based on expected value or linear utility. Bernoulli later developed the concept of diminishing marginal utility leading to logarithmic utility. According to this theory, the value of an item should be determined by its utility and not by its price. The price of a product depends only on the product itself and is the same for everyone, but the utility depends on the particular circumstances of the person making the assessment. Therefore, a small gain is more important for the poor than for the rich, even though both receive the same amount (Chukwudum, 2016).

EUT has been a widely accepted theory of decision making under risk. In the risky environment, each choice leads to a set of possible outcomes and the probability of each outcome is known. This theory recognizes that when choosing risky alternatives, people tend to maximize the expected utility of their choices, i.e., they decide on the option with the highest weighted sum of the utility of individual outcomes according to their probabilities (Mongin, 1997; Levy, 1992).

EUT uses the Bayesian technique when making decisions under uncertainty faced by individuals acting with the goal of rational optimization. Bayesian technique is based on the principle of choosing the option that will maximize the utility as a result of multiplying the probabilities of events and the utility to be obtained (Bostancı, 2003, p. 3).

According to this theory, a rational decision maker evaluates risky and uncertain prospects by their expected utility and selects the option with the highest value. Economic choices are modeled as consistent with material interests and the maximization of expected utility. In EUT, risk attitudes are not assumed a priori but represented by the curvature of the utility function; individuals may be risk-averse, risk-neutral, or risk-seeking (Mongin, 1997, p. 1).

EUT is based on the assumption that people will act rationally under all circumstances when making decisions, that individuals can accurately calculate the probabilities of the options they face before making a decision, and that they will choose the option that maximizes their utility as a result of this calculation. This time, rational behavior has been transformed into a full and complete calculation, advocating the idea that all individuals can make this calculation correctly (Sefil Tanrısever and Çilingiroğlu, 2011, p. 250).

EUT contains various assumptions. If we list these assumptions as items, they are as follows (Bailey, 2005; Bostancı, 2003):

- Man is a rational being. A rational person acts in line with his/her own interests, free from emotions, in order to maximize his/her interests while making decisions.
- Rational people do not show partiality to any preference in case of uncertainty. They start by using Bayes' Theorem in their choice of preference.
- Decision makers act according to the opportunities to maximize utility in uncertain situations. Between any two options, the one with the higher utility is preferred.
- If option A is preferred to option B and option B is preferred to option C, option A will be preferred to option C in all circumstances.

We can explain the basic thesis of EUT in explaining the rational behavior of the consumer as follows: In situations of uncertainty involving more than one option, the individual can simultaneously measure the probabilities of the options he faces and the utility he will obtain if he chooses these probabilities independently of each other. In order to make these measurements, what is needed is a mathematical expression of the two probabilities between which the individual will be indifferent in the face of two options, one of which has a certain outcome and the other of which is uncertain.

EUT continues to be the dominant approach for modeling decision making under risk. It has been used as a primary decision-making tool since World War II, to generate economic and financial forecasts, to identify biases in management science, and to describe various aspects of psychology. Furthermore, expected utility is a widely used economic method for decision making in public policy (List and Haigh, 2005).

2.3.3. Prospect Theory

This theory was proposed by Kahneman and Tversky (1979) as a criticism of traditional finance theories and especially of the EUT. Prospect Theory (PT) examined the shortcomings of EUT and systematized it as an alternative model. The main divergence between the two theories stems from the way they deal with human beings (Kahneman, 2017). As a matter of fact, PT basically argues that individuals can act intuitively and emotionally against the assumption of EUT that individuals will act rationally when making decisions under uncertainty (Ruben & Dumludağ, 2015, p. 44).

Expected Utility Theory (EUT) is the standard framework for decision-making under risk; it serves as a normative model of rational choice and has descriptive applications in economics. Outcomes are evaluated by weighting utilities with their probabilities. Within EUT, risk aversion over gains follows from decreasing marginal utility, since a concave utility function implies a preference for safer prospects with the same expected value. Kahneman and Tversky (1979, p. 263), however, document systematic asymmetries relative to a reference point: individuals tend to be risk-averse for gains but risk-seeking for losses, and losses carry greater weight than equivalent gains. Because EUT evaluates final wealth with a single concave utility function, it does not generate this pattern. Prospect Theory formalizes these features by defining value over gains and losses around a reference point and by using a value function that is concave over gains, convex over losses, and steeper for losses.

According to PT, people in a loss or gain situation have different preferences. Expressing the same problem in different ways can lead to changes in people's preferences depending on the problem. This can lead to risk acceptance or risk aversion, which is not always rational (Tversky & Kahneman, 1974, p. 3). According to the theory, individuals generally give more weight to losses than to comparable gains and generally tend to be risk averse in terms of gains, while they tend to accept risk in terms of losses (Levy, 1992).

Kahneman and Tversky refer to this as the reflexivity effect. Reflection effect can be defined as the decrease in sensitivity to changes in assets as the individual moves away from the point of gain or loss. According to the theory, the pain that individuals feel in loss is much greater than the satisfaction of gain. Therefore, individuals may turn down gains by making irrational decisions due to the predominance of risk aversion in their economic decisions (Şentürk and Fındık, 2014).

PT uses a descriptive decision analysis method in decision-making situations against risky financial events and examines how investors make decisions in risk environments (Barberis et al., 1998). Therefore, instead of relying on the assumption of rational behavior in people's decision-making processes, PT argues that people may have false beliefs or preferences and therefore may not be fully rational (Ritter, 2003).

Following Kahneman and Tversky, Richard Thaler developed the concept of mental accounting and showed how it shapes consumer and investor choices. Early contributions on consumer choice and behavioral regularities provided the foundations (Thaler, 1980), the core formulation of mental accounting was set out in Mental Accounting and Consumer Choice (Thaler, 1985), and a comprehensive synthesis followed in Mental Accounting Matters (Thaler, 1999). Thaler also popularized policy applications through the "nudge" framework (Thaler & Sunstein, 2008), and his broader integration of psychology into economics was recognized by the 2017 Nobel Prize in Economic Sciences.

2.4. Investor Biases in Behavioral Finance

2.4.1. Cognitive Biases

Cognitive biases encompass errors in processing accumulated knowledge and experiences during decision-making processes. These biases emerge when there is a discrepancy between attitudes and behaviors, influencing individuals' decision-making situations. Such cognitive biases are not limited to uneducated or irrational individuals; they can be observed with remarkable consistency and continuity across all segments of society (Hanson & Kyser, 1999, p. 635). These biases will be examined under the following headings: overconfidence, excessive optimism, conservatism, representativeness, cognitive dissonance, anchoring, and the gambler's fallacy.

2.4.1.1. Overconfidence bias

Overconfidence bias refers to investors' bias to overestimate their own knowledge and abilities regarding market conditions, which can lead to risky investment decisions. This behavioral bias has significant implications in finance and investing, contributing to excessive trading behaviors, suboptimal investment choices, and distorted market perceptions (Loppies et al., 2022).

At its core, overconfidence bias operates through several psychological mechanisms. First, it leads individuals to overestimate their predictive abilities and investment acumen. Research shows that overconfident investors trade more actively, believing they can outperform the market with their decisions (Pikulina et al., 2017). This increased trading frequency typically results in higher transaction costs and lower overall returns, as overconfident investors fail to adequately account for the risks inherent in their choices (Singh et al., 2024).

Overconfidence can be exacerbated by factors such as financial literacy and demographic characteristics. While financial literacy may serve as a mitigating factor, studies indicate that even highly financially literate individuals—particularly women—still exhibit significant overconfidence in investment decisions (Adil et al., 2021). This paradox creates a situation where investors, despite having better knowledge, may still fall victim to overconfidence biases that distort their decisions.

The interaction between overconfidence and external influences, such as market behavior and peer decisions, further complicates investment outcomes. For instance, herd behavior can amplify overconfidence, as investors may take more aggressive investment positions without due diligence when encouraged by their peers' actions (Kaban & Linata, 2024). This bias can contribute to market volatility and potentially lead to significant mispricing.

Empirical studies demonstrate a clear positive correlation between overconfidence and investor decisions, with this bias having detrimental effects on portfolio performance. Investors often hold losing stocks longer in hopes of a recovery while prematurely selling winning stocks to lock in gains—a phenomenon related to the disposition effect (Kansal & Singh, 2018). Such behaviors reveal a mismatch between perceived and actual market competence, complicating portfolio management efforts.

The effects of overconfidence are not limited to retail investors; they extend to corporate finance, where overconfident managers may take excessive risks in investment strategies, leading to suboptimal firm performance (Tekin, 2018). The reinforcement of such behavioral biases among decision-makers has broad implications for corporate strategy and overall market health.

2.4.1.2. Optimism bias

People are generally taught to believe in their abilities and traits, as well as to remain hopeful and maintain a positive outlook on life. This optimistic approach carries over into decision-making processes. Kahneman and Riepe (1998) define optimists as individuals who underestimate the probability of adverse outcomes beyond their control. Building on this idea, excessive optimism can manifest as a belief that one's choices will turn out favorably and that one's predictive ability is higher than warranted; overly optimistic individuals both exaggerate positive outcomes and downplay negative ones (Shefrin, 2006)..

Generally, when making decisions, individuals tend to be positive about the outcome, and investment decisions are no exception. While these thoughts help investors take on the inherent risks in financial decision-making, they do not eliminate them. This optimistic approach makes investors overconfident in their abilities and prone to overestimating their judgment (Joo & Durr, 2017).

Optimism is an individual difference trait reflecting the extent to which people hold generalized positive expectations about their future. Higher levels of optimism are prospectively associated with greater subjective well-being during times of distress or difficulty. The optimism bias is linked to beliefs in predicting events and the bias to attribute good outcomes to oneself while blaming bad outcomes on chance. This trait leads investors to believe their portfolio returns will outperform market averages and prompts them to invest more heavily in individual stocks (Oran & Perek, 2013).

Optimism, one of the biases that expose individuals to probabilistic risks with asymmetric effects depending on their goals, involves overestimating favorable information when evaluating events. Consequently, investors may perceive risks differently than they are, leaving them vulnerable to adverse situations. For example, a financial analyst's optimistic assessment of affiliated brokerage firms may encourage investors following that analyst to buy (Karakuş, 2022).

2.4.1.3. Conservatism bias

Individuals often interpret information in ways that align with their pre-existing beliefs rather than updating objectively. This tendency pairs with conservatism bias—the inclination to underreact to new evidence and rely on prior views when incorporating information or forming predictions. In practice, individuals weight initial beliefs too heavily and adjust too slowly as fresh data arrive (Pompian, 2012).

When economic factors constantly change during investment processes, investors encounter numerous ideas and new information. Conservatism bias emerges when investors hesitate to modify previous decisions or respond only partially to new developments in existing processes or emerging situations (Barberis et al., 1998, pp. 9-10).

In literature, "conservatism" is sometimes perceived merely as religiosity. However, in finance, it describes traditionalism and resistance to change. Behavioral finance views this as a negative behavioral pattern. People should remain open to change when necessary and be willing to revise previous decisions in light of new information. When conservatism prevails, individuals tend to avoid changes and cling rigidly to tradition. Letting go of existing knowledge becomes difficult, though they may eventually recognize the need for reassessment when significant events affect their investments.

Conservatism means individuals are slow to change their beliefs when confronted with new information and may fail to fully account for the implications of earnings surprises (Byrne & Brooks, 2008). People often perceive deviation from the norm as risky and prefer not to stray from their long-held decisions and preferences. When new opportunities arise, their sceptical approach and reluctance to abandon familiar patterns may cause them to miss potential advantages. For them, historically accepted information carries more weight as the most valid truth.

Due to conservatism bias, investors tend to react rigidly when receiving new information about subjects they already know. They typically cling to prior positive information while ignoring negative updates on the same topic. Because of conservatism, investors respond to market changes much more slowly compared to representativeness bias. While investors overreact to new information under representativeness bias, they underreact under conservatism bias (Jain & Kesari, 2019).

2.4.1.4. Representativeness bias

Representativeness bias is one of the important mental shortcuts affecting financial investment decisions. This bias causes investors to make mistakes in predicting financial asset prices and market movements. It causes investors to ignore real risks and probabilities when trying to predict future situations based on historical data or specific samples.

This bias, first introduced by Kahneman and Tversky (1972), suggests that things with broadly similar characteristics are similar to each other. In other words, it is to assume that a phenomenon that stands out with some characteristics is also good in other areas. For example, investors often see good companies as companies they can invest in. They consider a company to be a good investment choice because it produces quality products, is well-managed and has a continuous growth trend. They do not consider the possibility that the performance of a company that looks good now will decline in the future. They may be mistaken in their investment preferences by emphasizing the appearance of the company rather than statistical and financial data (Gümüş et al., 2013, p. 77).

According to Tversky and Kahneman (1974), representativeness bias involves making judgments based on stereotypes rather than the underlying characteristics of decision making. When individuals are under the influence of representativeness bias, events are classified as representative of a class well known to them. The consequence of such a bias is that probability estimates are made in a way that overemphasizes the importance of categorization without paying sufficient attention to the evidence about the underlying probabilities.

Representativeness bias is a cognitive bias in which investment decisions are made based on mental stereotypes without regard to actual and known experience. The success of these investors is likely to be repeated in the future. Due to representativeness bias in the investment decision-making process, investors may mistakenly attribute a company's positive characteristics to a good investment, lead to cognitive errors such as all positive characteristics of a company make it a smart investment, and view recent past performance as an indicator of future returns (Ayaa et al. 2022).

In representativeness bias, investors tend to make decisions based on surface features rather than underlying probabilities (Byrne and Brooks, 2008). Representativeness bias refers to the error of making evaluations based on one's own knowledge, experience, observations, etc. instead of calculating the probability of a situation or event (Karakuş, 2022).

It is seen that there are two different generalizations explained by the law of small numbers in the bias to surrender. In the first case, it is believed that large samples in a population reflect all the characteristics of the population and that the conditions valid for large samples are valid for small samples (Sefil Tanrısever & Çilingiroğlu, 2011). For example, based on the information that mutual funds have had good returns in the last two years, a prediction is made that the future returns of a fund will perform similarly. In the other case, it is the belief that a small sample has the general characteristics of the population and generalizes to a large sample (Barberis & Thaler, 2003; Demirel & Yelkikalan, 2022). For example, looking at a fund with a good return and believing that the return of all funds will also be high.

Another misconception is the belief that the current situation will continue. Investors tend to assume that recent events will continue in the future and prefer to buy emerging stocks rather than stocks that have recently performed poorly (Arslan & Boztosun, 2022, p. 32).

2.4.1.5. Mental accounting bias

Mental accounting, first introduced by Thaler (1980), is defined as the bias of people to codify, classify, and value economic outcomes by grouping their assets into a large number of immutable mental accounts. In other words, it is a bias that expresses how individuals allocate all their money in their minds according to personal factors.

The prejudice that investors open new accounts in their minds for each decision they make and think about their investment activities independently of each other can be referred to as mental accounting. Due to the lack of interaction between these accounts, investors' perceptions of adequacy decrease, risks are not recognized or are misperceived, and timing mistakes are made (Şamandar & Gömlekçi, 2019).

2.4.1.6. Cognitive dissonance

Cognitive dissonance, which was first introduced by Leon Festinger (1965), is defined as a situation where an individual's beliefs or knowledge do not match with each other. According to this bias, people define their behaviors and thoughts according to their past values. These can be beliefs, attitudes and needs. All these values or environmental factors we acquire over time shape our personality.

Cognitive dissonance is a phenomenon encountered in both psychology and behavioral finance. It refers to the contradiction/inconsistency between an individual's attitudes, feelings, thoughts, beliefs and behaviors. For example, if a person continues to smoke even though he/she believes that smoking is harmful to his/her health, then cognitive dissonance arises (Kocabey, 2022).

Cognitive dissonance, the bias to adjust beliefs to justify past actions, is a psychological phenomenon called cognitive dissonance. Individuals are disturbed by contradictory cognitive elements, such as the inconsistency between empirical evidence and past choices, and they change their beliefs to reduce this discomfort. Since they tend to ignore new data that invalidate known beliefs and choices, the mental conflict they face when they realize that their beliefs and suspicions are not correct leads them out of line and to irrational investment decisions (Virigineni and Rao, 2017).

Since people see themselves as smart, capable of making the best decisions, when faced with a negative situation, they tend to ignore and reject information that contradicts their own thoughts and decisions, which is called cognitive dissonance. When investors fall into cognitive dissonance, they hold their securities at a loss even though they have to sell them due to the psychological discomfort of admitting that they have made the wrong decision. Furthermore, they continue to invest in shares whose value falls below normal levels. This may lead them to avoid information that contradicts their previous decisions until they drown in it and believe that they are different this time in their new investments (Elmas, 2014).

When individual investors encounter information that contradicts their own knowledge, they may face financial losses because they pretend not to see it. This is due to the low impact level of information that contradicts the mental data record and the perception of positive data as acceptable (Ertuğrul Ayrancı, 2020).

There are three possibilities for reducing cognitive conflict: first, if two perceptions are incompatible, one can be changed to fit the other. Second, if two perceptions cause a significant conflict, the degree of conflict can be reduced by creating one or more congruent perceptions, and finally, third, another way to reduce conflict is to increase identity awareness (Festinger, 1965).

2.4.1.7. Anchoring bias

This bias, articulated by Kahneman and Tversky (1974), assumes that the human brain, in the face of complex problems, sets a starting point and then slowly adjusts the reference point in line with the information added. In other words, people base their forecasts on the reference point they have set. Past knowledge, experience and observations are effective tools in decision making. Accordingly, people make decisions based on these experiences. These situations are explained by the anchoring technique's assumption that forecasts are made based on a reference point (Aktaş, 2021)

The anchoring bias explains the fact that when creating quantitative estimates of uncertain quantities, evaluations based on some initial values are often inadequate. In financial markets, investors can make buying and selling decisions based on a reference point. Therefore, anchoring bias can be defined as an investor's acceptance of a particular price of a security as a reference point and making decisions based on that reference point (Full Cen et al., 2013).

2.4.1.8. Gambler's fallacy

It is a bias that people often fall into when predicting successive outcomes in random events. The similarity between random processes is incorrectly interpreted by an investor as a predictive relationship between them (Rakesh, 2013). Under this bias, for example, after several consecutive heads in fair coin tosses, one may falsely believe that tails has become more likely; yet successive trials are independent and the probability remains unchanged. Classic work by Tversky and Kahneman (1974) attributes this error to the representativeness heuristic and the related belief in the “law of small numbers,” whereby people expect short sequences to mirror long-run frequencies.

The gambler's fallacy is the fallacy that an event can be predicted by chance or coincidence. These fallacies are common everywhere in life in the possibility of independent events. This bias arises when investors predict the outcome of market returns to be good or bad (Ceylan and Korkmaz, 2022).

Events from the past are often a reference for investors to determine their investment decisions. In this fallacy, there is a bias for investors to make predictions based on the results of previous events in independent and random events (Karan, 2018). When an investor has achieved returns several times in the previous period, they tend to reduce their level of investment for the future. During an uptrend, investors with a gambler's bias will tend to avoid buying stocks that have previously experienced price increases, assuming a larger price decline in future periods. Conversely, when a downtrend occurs, investors with a gambler's bias will tend to buy stocks that have previously experienced price declines, assuming a larger price increase in future periods (Wijayanti et al., 2019).

Gambler's fallacy is defined as the erroneous belief that people develop for independent events. They believe that events are interconnected. For example, when playing a lottery, people would never bet on a number they bet on in a previous event. The gambler's fallacy is thought to be caused by the representativeness bias or the Law of Small Numbers. While the gambler's fallacy assumes that the event that has already occurred will not repeat, the hot hand fallacy thinks that it will (Asad et al., 2018).

2.4.2. Emotional Biases

Emotion is a psychological change in an individual's mood caused by himself/herself, his/her experiences or his/her environment. Emotions, which are an important factor in many behavioral theories, were not previously considered as a factor affecting the decision-making process, but with the intensification of behavioral finance research, they have become one of the important factors affecting the decision-making process. In this context, emotional tendencies will be analyzed under the following headings: loss aversion, predisposition bias, ownership effect and sadness bias.

2.4.2.1. Loss aversion

Kahneman and Tversky (1979) proposed that loss aversion is the bias of individuals to be more sensitive to loss. In other words, individuals tend to be more loss averse than gain averse. According to expectancy theory, loss aversion behavior refers to people's bias to prefer avoiding loss to gain. This bias depends on past gains and losses. There is an emotional connection between investors and their current assets. It is seen that people prefer not to lose their current assets rather than choosing what they can gain. The weight given to losses is higher than that given to gains of the same amount. This is because people are averse to losses as they appear larger than gains (Prosad et al., 2015, p. 8).

The risks an individual is willing to take are not only context-dependent but also asymmetric. In other words, because individuals are loss averse, losses appear larger than gains (Kahneman and Tversky, 1979, p. 279) and losses are more painful than equal gains (McDermott, 2004, p. 298). Loss aversion is an important variable in terms of risk-taking bias. People shy away from change and thus reinforce the status quo. Specifically, individuals have a strong bias to stay with the status quo because the disadvantages of leaving it seem greater than the advantages (Kahneman et al., 1991, p. 198).

Loss aversion is a common bias among individual investors. Limited attention and mental capacity may lead them to confuse the unpleasantness of experiencing economic loss with the unpleasantness of recognizing it. This behavior is related to the concept of loss aversion, where people are concerned with gains and losses measured against an arbitrary reference point. When the stocks that investors decide to sell outperform the stocks they hold, it is contrary to the principles of investing, but investors do not tend to dispose of stocks where profits start to accrue and to cut losses by holding stocks where losses start to accrue (Daniel et al., 2002).

People value returns based on whether there is a gain or loss compared to their status quo position. This is related to the idea that individuals adjust to a level of income to which they are accustomed, so that subjective well-being is associated with changes in income rather than the level of income. Indeed, people are about twice as reluctant to suffer losses in order to receive equal gains. Therefore, investors may tend to be loss averse by avoiding losses rather than seeking gains (Virigineni and Rao, 2017).

2.4.2.2. Divestment bias

Divestment bias is defined as the bias of investors to sell winning investments too early and hold losing investments for too long (Bernard et al., 2022). At the core of the propensity to divest is Kahneman and Tversky's concept of loss aversion, which is prominent in IT. According to this theory, investors are psychologically more affected by losses than gains of the same magnitude, which leads them to take excessive risks on their lost investments in the hope of recovering their losses, while in case of profits, they tend not to take risks and tend to divest quickly to realize their profits (Da Silva et al., 2008).

This behavior is further complicated by mental accounting, where individuals categorize their investments in separate accounts based on performance, treat realized gains differently from unrealized losses, which can distort rational reasoning (Ammann et al., 2011).

Empirical evidence supports the existence of the divestment effect in various market contexts and studies show its prevalence among retail and institutional investors worldwide (Bernard et al., 2022; Cici, 2012). Cici (2012) argues that mutual funds often exhibit this bias by holding underperforming assets for longer than necessary, a bias that harms overall portfolio performance.

Moreover, research shows that the divestment effect is not uniform across different investor demographics and can be influenced by factors such as gender, age, and investment experience (Breitmayer et al., 2019; Zahera and Bansal, 2019). For example, studies have found that less knowledgeable investors are more prone to the bias effect than their more knowledgeable counterparts. This suggests that education and experience in the investment world may reduce this behavioral bias (Dhar & Zhu, 2006; Feng & Seasholes, 2005).

Interestingly, this phenomenon is also linked to market dynamics, where group behavior associated with the propensity effect can lead to stock price momentum. Investors' collective selling of winners too early can create upward price pressures, while their refusal to sell losers can cause downward pressures, leading to market mispricing and contributing to greater volatility (Muermann & Wise, 2006).

2.4.2.3. Avoiding regret

Regret is the sadness that individuals feel when a decision they have made ends in an undesirable way. Individuals who tend to avoid regret are often afraid and hesitant about the consequences of the decisions they take or have taken. In other words, the possibility of experiencing regret can affect individuals' decisions. Regret avoidance refers to the fact that people make their decisions accordingly in order not to regret the decisions they will take or the decisions they will make later. Therefore, regret avoidance argues that people tend to reduce risk or prefer safer options by focusing on the negative consequences of a potential decision. Postponing decision-making due to the fear of regret is also considered as a reflection of regret avoidance (Seiler et al., 2008).

In finance, two types of regret can be mentioned. One of them is the regret arising from doing something, while the other is the regret arising from not doing something. However, research on the subject has shown that the regret people feel after doing something causes a greater sadness than the regret caused by not doing something (Virigineni and Rao, 2017; Zeelenberg et al., 2002, p. 314).

Investors tend to sell their winning investments earlier in order to be proud of the results of their investment decisions and hold their losing investments for a longer period of time in order not to feel regret (Gül, 2017). In addition, consistent with the sunk-cost effect, some investors add to losing positions to "average down" rather than realize a loss, thereby escalating commitment to prior choices (Arkes & Blumer, 1985).

2.4.3. Social Biases

According to the behavioral finance approach, the social environment has a significant impact on human decisions. Social factors affecting investor behavior include the influence of family members, friends, and colleagues on whether or not to buy a particular financial instrument. Investors behave irrationally under the influence of these social factors (Asad et al., 2018). Herd psychology, which falls under the category of social tendencies, will be analyzed below.

2.4.3.1. Herding Behaviour

Herding is an important behavioral finance phenomenon characterized by the bias of individuals to imitate the actions or decisions of a larger group, often causing them to disregard their own knowledge or analytical skills. This behavior can affect investment decisions, market dynamics, and contribute to phenomena such as price volatility and bubbles (Shahani & Ahmed, 2023).

The essence of herding behavior is based on psychological and social factors. Investors may succumb to herding behavior for various reasons, such as reducing perceived risk, gaining social approval, or avoiding feelings of regret if their independent decisions turn out to be wrong. Research has shown that when individuals perceive risk to be high, they are more likely to follow the crowd, believing that collective action will yield better outcomes (Wibowo et al., 2023). For example, studies suggest that risk perception plays a mediating role between herding behaviour and investment decisions, especially among younger investors who may have less experience (Kaban and Linata, 2024).

Empirical evidence suggests that herd behavior can significantly influence investment decisions. In equity markets, herding behavior can cause price movements that do not reflect underlying fundamentals, often leading to inflated asset prices during bubbles and sharp declines during crashes (Lasantha & Kumara, 2021). A study by Qasim et al. (2019) highlighted how both herd psychology and overconfidence biases profoundly affect decision-making among investors in Pakistan and showed that these biases can lead to collective irrational behavior that deviates from sound investment principles.

Moreover, the impact of herding is not uniform across demographic groups; it can vary significantly depending on factors such as financial literacy and risk tolerance (Apochi et al., 2024). For example, some research suggests that financial literacy may moderate the impact of herding on investment decisions, suggesting that more informed investors may be better equipped to resist the pull of social influences (Adil et al., 2021). In contrast, less informed investors may be more prone to herding, which may lead them to act on impulse rather than informed analysis (Apochi et al., 2024).

2.5. Decision Tree Methods and Previous Studies in Behavioral Bias Detection

2.5.1. Studies on Behavioral Bias Detection in Financial Markets

The detection of psychological biases that influence investor behavior in financial markets is one of the fundamental research areas in behavioral finance literature. Behavioral biases can significantly disrupt investment decision-making processes and prevent investors from making rational decisions (Barber & Odean, 2001). Therefore, it is important for investors to identify and understand their behavioral biases in order to improve their decision-making abilities (Moore & Healy, 2008).

2.5.2. Studies on the Detection of Behavioral Biases Using Decision Tree Methods

2.5.2.1. Studies on herd behavior

Herd behavior occurs when investors disregard their own information and follow the behavior of others. Bikhchandani and Sharma (2001) demonstrated that herd behavior can significantly influence price formation in financial markets. Kukacka and Barunik (2012) showed that herd behavior can be numerically defined from transaction data and classified using decision trees. Herd behavior can be identified through transaction pattern analysis, especially during periods of high market volatility.

2.5.2.2. Studies on overconfidence

Overconfidence refers to investors' misguided excessive beliefs about their knowledge and abilities. Barber and Odean (2001) demonstrated that overconfident investors tend to trade more frequently, which typically results in lower returns. Moore and Healy (2008) emphasized that overconfidence distorts investment decisions and can be measured through high trading frequency and low performance patterns in transaction data.

2.5.2.3. Studies Related to Loss Aversion

Loss aversion is the tendency to feel losses more intensely than equivalent gains (Kahneman & Tversky, 1979). Liu et al. (2014) found that investors avoid realizing losses in large-scale transaction data and that this tendency can be identified using decision trees. The tendency to hold onto losing positions for an extended period is a typical indicator of loss aversion behavior.

2.5.3. Integrated Studies

Machine learning techniques such as decision trees enable the systematic identification of herd behavior, overconfidence, and loss aversion from transaction data. Noti et al. (2016) emphasize that behavioral decision-making processes can be successfully represented by machine learning models. This integration can help to better understand investor behavior and develop intervention strategies.

2.5.4. The Role of Financial Literacy

Financial literacy is an important regulatory factor in reducing the impact of behavioral biases on investors. Investors with high levels of financial literacy may be less susceptible to biases such as herd behavior and overconfidence (Moore & Healy, 2008). Therefore, financial literacy education can play a critical role in improving investment decisions.

2.5.5. Anomaly Detection Approaches

Anomaly detection in transaction data can reveal the effects of behavioral biases. Kukacka and Barunik (2012) noted that unusual movements in financial transactions are associated with biases such as herd behavior or overconfidence. Decision trees and anomaly detection methods enable a more in-depth analysis of investor behavior.

2.5.6. Machine Learning and Clustering Techniques

The use of machine learning and clustering techniques is widespread in the detection of behavioral biases. Noti et al. (2016) successfully predicted human behavior and identified investor groups using hybrid machine learning approaches. These methods are important for creating investor behavioral profiles and developing risk management strategies.

2.5.7. Data Set Applications

This study utilizes the FAR-Trans data set. Sanz-Cruzado, Droukas, and McCreadie (2024) have presented this dataset as a comprehensive resource for examining investor behavior. The dataset contains transaction data from European investors between 2018 and 2022 and is suitable for examining behavioral finance phenomena such as herd behavior, overconfidence, and loss aversion.

3. METHODOLOGY

This study employed a quantitative research approach to identify and classify behavioral biases among individual investors using machine learning techniques. The research model was built upon three fundamental biases encountered in behavioral finance theory: loss aversion, herding, and overconfidence. A combination of descriptive and predictive analytical methods was utilized. The initial phase involved calculating metrics representing behavioral biases from investor transaction data. In the subsequent phase, these metrics were normalized, and each investor was assigned to the behavioral group corresponding to their highest normalized score. Finally, a classification model was developed using a decision tree algorithm (Breiman et al., 1984).

3.1. Data

3.1.1. Data Source and Scope

The study utilized the FAR-Trans (Financial Asset Return Transactions) dataset, provided by the University of Glasgow. This comprehensive dataset encompasses 388,049 unique transaction records (totaling 154,103 unique trades) belonging to 29,090 individual investors between 2018 and 2022. The dataset is structured into four primary files:

`transactions.csv`, `customer_information.csv`, `asset_information.csv`, and `close_prices.csv`.

Within this dataset, there are 806 unique financial assets, with 321 having at least one recorded transaction. The dataset also includes 228,913 unique buy transactions and 159,136 unique sell transactions.

3.1.2. Data Pre-processing

The raw data underwent several pre-processing steps to ensure data quality and suitability for analysis:

- **Data Cleaning:** Date and time columns were converted to appropriate formats to facilitate accurate analysis.
- **Outlier Treatment:** To mitigate the influence of extreme values in the calculated behavioral metrics, the Winsorization technique was applied (Barnett & Lewis, 1994). Values falling outside the 5th and 95th percentiles were capped at their respective threshold values (Tukey, 1962).
- **Inclusion in Analysis:** Following the data cleaning procedures, 5,818 customers were randomly separated for the test set, ensuring an independent evaluation of the model's performance.

3.1.3 Calculation of Behavioral Metrics

We compute three behavioral metrics per investor: loss aversion, herding, and overconfidence. Each metric is defined in a way that is consistent with the behavioral-finance literature and directly implementable on our transaction-level data.

- Loss Aversion: Guided by prospect theory (Kahneman & Tversky, 1979) and the disposition-effect literature (Shefrin & Statman, 1985; Odean, 1998), we operationalize loss aversion via a holding-time proxy. Buy and sell transactions are matched on a first-in, first-out (FIFO) basis to form closed positions. For investor i, let H^{Loss}_i and H^{gain}_i denote the average holding period (in days) of positions closed at a loss and at a gain, respectively. The loss-aversion score is then defined as:

$$\mathrm{LA}_i = \frac{\overline{H}_i^{\mathrm{loss}}}{\overline{H}_i^{\mathrm{gain}} + c},$$

where c = 1 is a small smoothing constant to avoid division by zero when $H^{gain}_i \approx 0$. Higher values indicate a stronger tendency to hold losers longer than winners, which is the core behavioral regularity emphasized by the disposition-effect literature. We adopt this investor-level holding-time implementation because it maps transparently onto our microdata while remaining conceptually aligned with established definitions.

- Herding: Following the herding framework of Bikhchandani and Sharma (2001) and consistent with empirical approaches to collective behavior in markets (Christie & Huang, 1995), we construct a trade-popularity measure at the investor–month level. For each month t, we identify the Top-5 most traded assets by total transaction count in the market. For investor i, the herding ratio is:

$$H_{i,t} = \frac{\text{trades}_{i,t}(\text{Top5}_t)}{\text{trades}_{i,t}(\text{All})},$$

  i.e., the proportion of that investor's trades executed in the month's Top-5 assets. Larger values capture a stronger tilt toward contemporaneously popular instruments.

- Overconfidence: In line with common practice that links excessive trading to overconfidence (Barber & Odean, 2001; Odean, 1999), we use trading frequency as a proxy. For investor i, average monthly trading activity is computed as:

$$\text{TradesPerMonth}_i = \frac{1}{T}\sum_{t=1}^{T} \text{trades}_{i,t},$$

  where $\text{trades}_{i,t}$ is the number of executed trades by investor i in month t. Higher frequency indicates greater overconfidence in this behavioral interpretation.

The descriptive statistics for these metrics are reported in Table 4.1.1.

3.1.4. Error Definitions and Computation (Multi-class)

We evaluate the three-class decision-tree classifier (Loss Aversion, Herding, Overconfidence) using accuracy, precision, recall, F1-score, and the threshold-independent summary ROC–AUC. In addition, we report Type I and Type II error rates per class using a one-versus-rest reduction of the confusion matrix. For class c, letting

TPc, FPc, TNc, and FNc denote true positives, false positives, true negatives, and false negatives:

$$\mathrm{FPR}_c = \frac{\mathrm{FP}_c}{\mathrm{FP}_c + \mathrm{TN}_c} \quad (\text{Type I}), \qquad \mathrm{FNR}_c = \frac{\mathrm{FN}_c}{\mathrm{FN}_c + \mathrm{TP}_c} \quad (\text{Type II}).$$

Overall accuracy is: Accuracy = $\frac{\mathrm{TP+TN}}{\mathrm{N}}$. All quantities are computed from the cross-validated predictions aggregated across folds. Definitions follow standard practice in classification evaluation (Sokolova and Lapalme, 2009).

3.1.5 Exploratory Unsupervised Clustering and Rationale for Supervised Modeling

Before training the multi-class classifier, we explored whether the feature space exhibits natural groupings. We implemented k-means clustering, which minimizes within-cluster variance under a Euclidean metric (MacQueen, 1967; Lloyd, 1982), and DBSCAN, a density-based method that identifies core points and expands clusters from them (Ester et al. 1996). For k-means, we inspected the within-cluster sum-of-squares elbow and silhouette profiles to assess separation (Rousseeuw, 1987). For DBSCAN, we examined alternative settings for the neighborhood radius ε\varepsilonε and the min_samples parameter to evaluate the stability of assignments.

Across reasonable parameter choices, these exploratory runs did not reveal a stable cluster structure that aligned with our three target categories (Loss Aversion, Herding, Overconfidence). Given the availability of reliable investor-level labels and our prediction objective, we proceeded with a supervised decision-tree model. The unsupervised results were not used for model selection, feature construction, or evaluation beyond this preliminary check.

3.2 Hypothesis

The core hypotheses formulated for this study are as follows:

• H1: Individual investors can be significantly divided into distinct groups based on loss aversion, herding behavior, and overconfidence biases.

• H2: Investor groups formed based on behavioral biases exhibit statistically significant differences in their calculated metrics.

• H3: The decision tree algorithm can predict investors' behavioral bias groups with high accuracy.

3.3. Supervised Modeling

3.3.1. Group Assignment and Statistical Tests

For each investor, the three calculated behavioral metrics were standardized using z-score normalization (Hair et al., 2014). Subsequently, each investor was assigned to the behavioral group corresponding to the metric with their highest normalized value. This methodology aligns with the dominant bias detection approach utilized in previous studies, such as that by Ben-David et al. (2013). The resulting distribution of group assignments is presented in Table 4.2.1. To assess the statistical significance of differences between these groups, a Kruskal-Wallis H test was applied (Kruskal & Wallis, 1952), as the normality assumption was not met, making it suitable for non-parametric comparisons across multiple groups (Field, 2013). The test results indicated that inter-group differences were statistically significant for all metrics ($p<0.0001$).

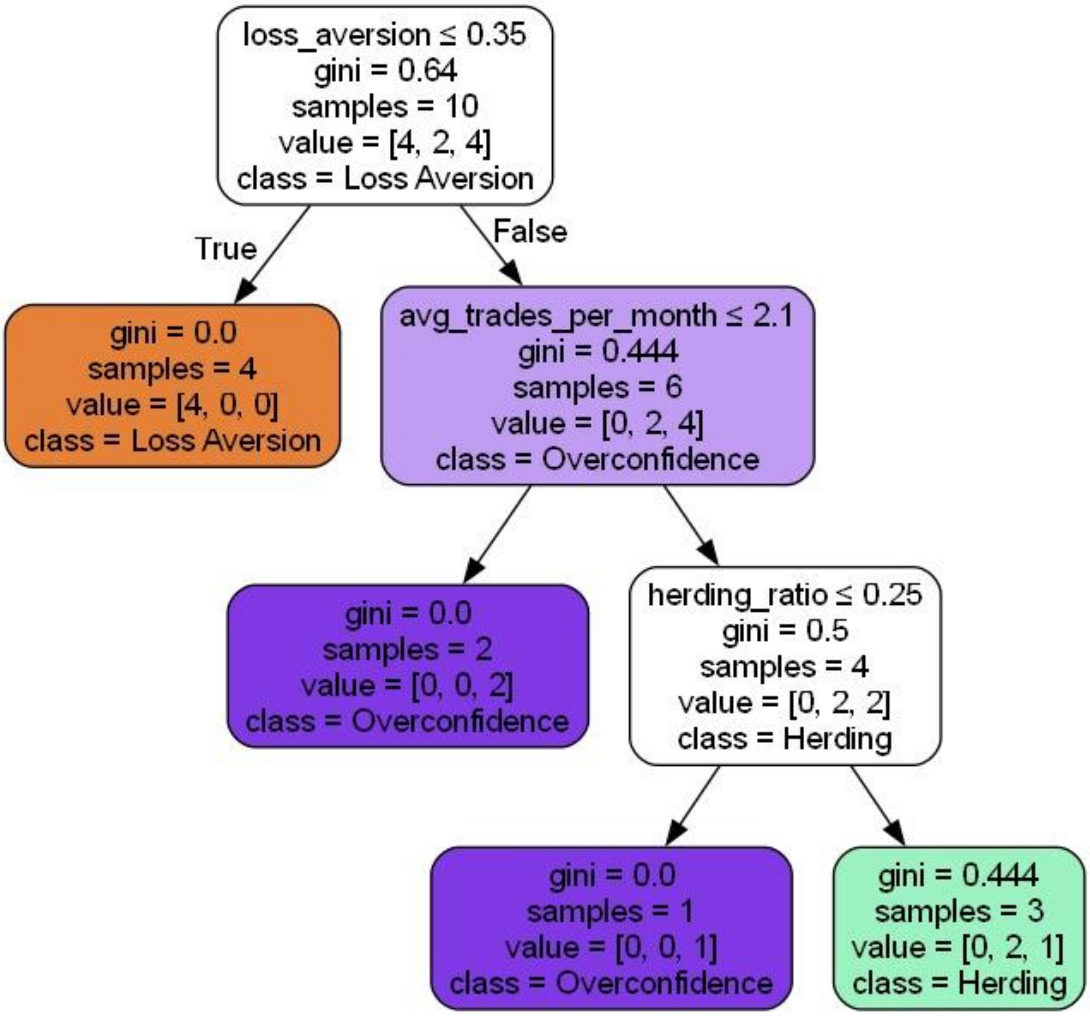
3.3.2. Model Development and Evaluation

A Decision Tree algorithm was selected for the classification problem due to its interpretability and effectiveness (Han et al., 2012). To minimize the risk of overfitting, the maximum depth of the model was limited to 3 (Hastie et al., 2009). The dataset was partitioned into training (80%) and testing (20%) sets using stratified random sampling (Kohavi, 1995), ensuring that each class was represented proportionally in both sets. Model performance was rigorously evaluated using a suite of widely accepted metrics:

- Accuracy
- Precision
- Recall (Sensitivity)
- F1-Score
- ROC-AUC values

These metrics are standard for evaluating classification model performance (Fawcett, 2006; Powers, 2011). Furthermore, to assess the model's generalization capability and robustness, 5-fold cross-validation was performed (Stone, 1974).

Figure 3.3.1 Decision Tree Visualization for Classification of Investor Behavioral Biases



4. FINDINGS

4.1. Descriptive Analysis of Behavioral Metrics

The empirical analysis commenced with a detailed examination of the three core behavioral metrics designed to capture distinct investment behaviors among individual investors. The study analyzed transaction data from 29,090 individual investors within the FAR-Trans dataset. After meticulous data cleaning procedures, 5,818 investors were set aside for the test set to ensure an unbiased evaluation of the classification model. Table 4.1.1 presents the descriptive statistics for the primary variables utilized in this study.

Table 4.1.1 Descriptive Analysis of Behavioral Metrics

Metric	Mean	Std Dev	Median	Minimum	Maximum
Loss Aversion	0.64	May.70	0.00	-0.11	368.25
Herding Ratio	0.35	0.32	0.15	0.05	0.94
Avg Trades Per Month	2.0	2.Oca	Oca.40	0.00	27.00

The descriptive statistics reveal that all three behavioral metrics exhibit right-skewed distributions with the presence of outliers, a finding consistent with behavioral finance literature suggesting heterogeneous investor behavior patterns (e.g., Kahneman & Tversky, 1979; Odean, 1998). The loss aversion metric demonstrates the highest variability (σ=5.70), indicating substantial differences in investors' tendencies to hold losing positions. The herding ratio shows moderate dispersion (σ=0.32), while trading frequency exhibits relatively consistent behavior across the sample (σ=2.1). A noteworthy observation is the median value of the loss aversion metric, which is 0.00. This suggests that a significant majority of investors did not realize losses during the analysis period, aligning with the disposition effect commonly observed in behavioral finance (Shefrin & Statman, 1985; Odean, 1998).

4.2. Behavioral Bias Group Formation and Distribution

Based on the dominance of each behavioral metric (i.e., the highest normalized score), investors were classified into three distinct behavioral bias groups. The classification process resulted in a relatively well-balanced distribution across the sample, as illustrated in Table 4.2.1.

Table 4.2.1 Behavioral Bias Group Formation and Distribution

Behavioral Group	Number of Individuals	Percentage (%)
Loss Aversion	8,879	30.8
Herding	8,666	30.0
Overconfidence	11,545	39.2
Total	29,09	100.0

The distribution demonstrates a relatively balanced representation across all three behavioral categories. Overconfidence emerged as the most prevalent bias, affecting 39.2% of investors, followed closely by loss aversion (30.8%) and herding behavior (30.0%). The significant proportion of investors in the overconfidence group suggests that a considerable portion of the investor population exhibits high trading frequency, which is often linked to overconfidence (Barber & Odean, 2001). This balanced distribution is highly advantageous for subsequent multi-class classification analysis, as it minimizes potential class imbalance issues that could adversely affect model performance (He & Garcia, 2009).

4.3. Comparative Analysis of Behavioral Metrics Across Groups

To validate the distinctiveness and inherent characteristics of each behavioral group, a comparative analysis was conducted to examine the mean values of the core metrics across the different bias categories. Table 4.3.1 presents the results of this comprehensive cross-group comparison.

Table 4.3.1 Comparative Analysis of Behavioral Metrics Across Groups

Bias Group	Loss Aversion (Mean ± SD)	Herding Ratio (Mean ± SD)	Avg Trades Per Month (Mean ± SD)
Loss Aversion	706.36 ± 573.57	0.07 ± 0.26	1.50 ± 1.71
Herding Behavior	0.07 ± 0.26	0.81 ± 0.30	1.13 ± 0.30
Overconfidence	0.38 ± 0.53	0.19 ± 0.53	2.49 ± 2.95

The results unequivocally demonstrate a clear and significant separation between groups based on their respective dominant behavioral metrics. The loss aversion group exhibits substantially higher loss aversion scores (Mean = 706.36 ± 573.57), indicating a strong tendency to hold onto losing positions for extended periods (Kahneman & Tversky, 1979; Shefrin & Statman, 1985). Concurrently, this group maintains remarkably low herding tendencies and a moderate trading frequency. In contrast, the herding group displays the highest herding ratio (Mean = 0.81 ± 0.30), signifying a pronounced inclination to follow the actions of others (Bikhchandani & Sharma, 2001), while showing minimal loss aversion behavior and the lowest trading frequency among the groups. Finally, the overconfidence group is distinctly characterized by the highest average trading frequency (Mean = 2.49 ± 2.95). This elevated trading activity is highly consistent with the overconfidence bias, where investors may believe their market timing or stock-picking abilities are superior, leading to more frequent transactions (Barber & Odean, 2001). Despite their high trading frequency, this group maintains moderate levels of the other behavioral metrics.

4.4. Statistical Significance Testing

To ensure the statistical validity of group differences, Kruskal-Wallis tests were conducted on the core behavioral metrics. Table 4.4.1 summarizes the results of these significance tests.

Figure 4.4.1 Boxplot of Behavioral Bias Groups

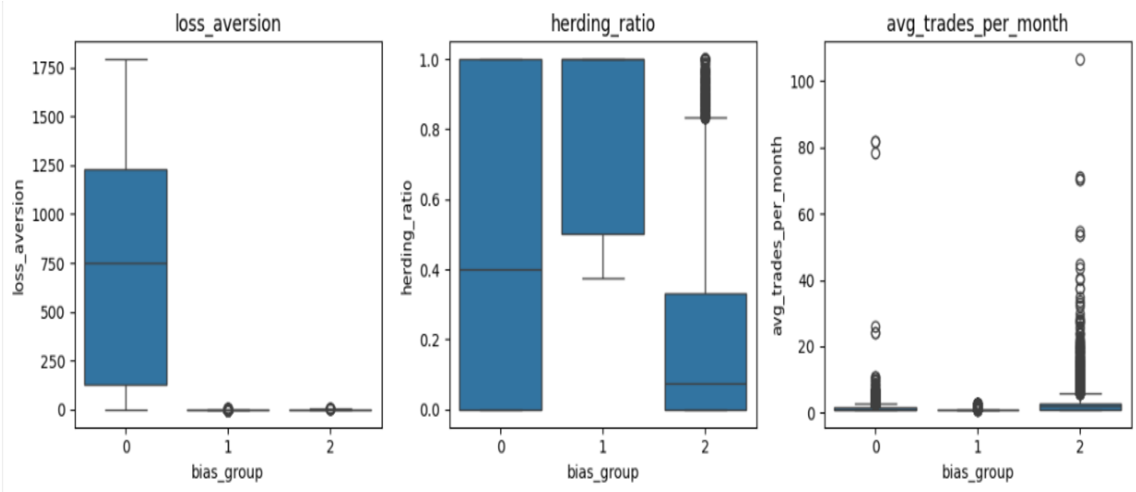


Table 4.4.1 Statistical Significance Test Results

Metric	H Statistics	p-value	Significany
Loss Aversion	15,847.23	< 0.0001	***
Herding Ratio	14,923.56	< 0.0001	***
Avg Trades Per Month	8,462.91	< 0.0001	***

Note: *** $p < 0.001$

The extremely low p-values ($p<0.0001$) for all metrics provide strong statistical evidence for significant differences between the three behavioral bias groups. These results confirm that the group classifications are statistically meaningful and not due to random variation in the data.

4.5. Classification Model Performance

A decision tree classifier was employed to evaluate the predictability and robustness of the behavioral bias groupings, serving as a critical step in validating the proposed framework. The model's performance on the unseen test set is presented in Table 4.5.1 providing a comprehensive overview of its predictive capabilities.

Table 4.5.1 Classification Model Performance

Behavioral Group	Precision	Recall	F1-Score	Support
Loss Aversion	0.99	0.96	0.97	1,776
Herding	0.97	0.94	0.96	1,733
Overconfidence	0.93	0.99	0.96	2,309
Macro Average	0.97	0.96	0.97	5,818
Weighted Average	0.96	0.96	0.96	5,818
Overall Accuracy			0.96	5,818

The classification model demonstrates exceptionally high performance across all evaluated metrics, achieving a remarkable overall accuracy of 96%. The precision scores, ranging from 0.93 to 0.99, indicate a minimal rate of false positive classifications, meaning that when the model predicts an investor belongs to a certain bias group, it is highly likely to be correct. Similarly, the recall scores, ranging from 0.94 to 0.99, demonstrate the model's excellent ability to correctly identify instances of each behavioral bias, minimizing false negatives. The F1-scores, which provide a balanced measure that considers both precision and recall, consistently exceed 0.96 across all groups, further highlighting the model's robustness and effectiveness in accurately classifying investors into their respective behavioral bias categories. These results are indicative of a well-performing classification model (Sokolova & Lapalme, 2009).

4.6. Model Validation and Reliability Assessment

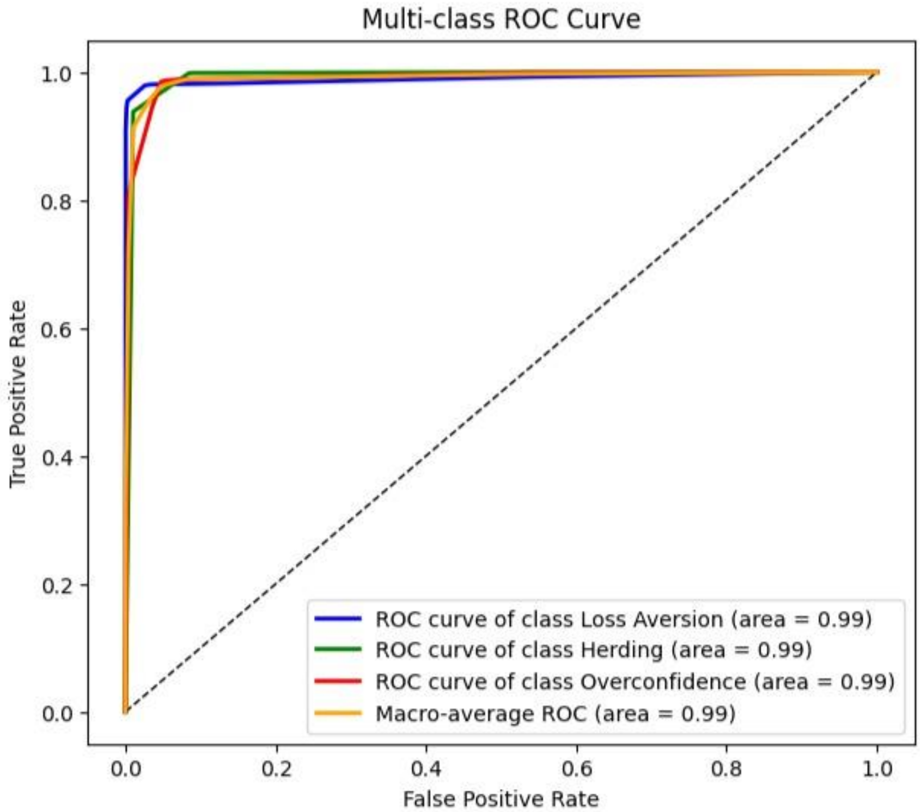
To rigorously assess the stability and generalizability of the classification model, a 5-fold cross-validation technique was performed (Kohavi, 1995). This method involves partitioning the dataset into five equal subsets, training the model on four subsets, and validating on the remaining one, iterating this process five times. Table 4.6.1 presents the cross-validation results.

Table 4.6.1 Model Validation and Reliability Assessment

Fold	Accuracy Score
1	0.9563
2	0.9608
3	0.9625
4	0.9644
5	0.9622
Mean	0.9613
Standard Deviation	0.0027

The cross-validation results demonstrate remarkable consistency across all folds, with individual accuracy scores ranging from 0.9563 to 0.9644. The low standard deviation of 0.0027 indicates minimal variance in model performance across different data subsets, strongly suggesting high reliability and excellent generalizability of the classification approach (James et al., 2013). This consistency across multiple validation sets provides strong evidence that the model is not merely performing well on a specific data partition but is robust and capable of accurately classifying new, unseen investor data.

Figure 4.6.1 Multi-Class ROC Curve



4.7. Discriminatory Power Analysis

The discriminatory power of the classification model, which refers to its ability to distinguish between different classes, was further rigorously evaluated using ROC curve analysis for each behavioral bias group (Fawcett, 2006). The Area Under the Curve (AUC) scores derived from this analysis are summarized in Table 4.7.1.

Table 4.7.1 Discriminatory Power Analysis

Behavioral Group	ROC-AUC Score
Loss Aversion	0.99
Herding	0.99
Overconfidence	0.99
Macro Average	0.99

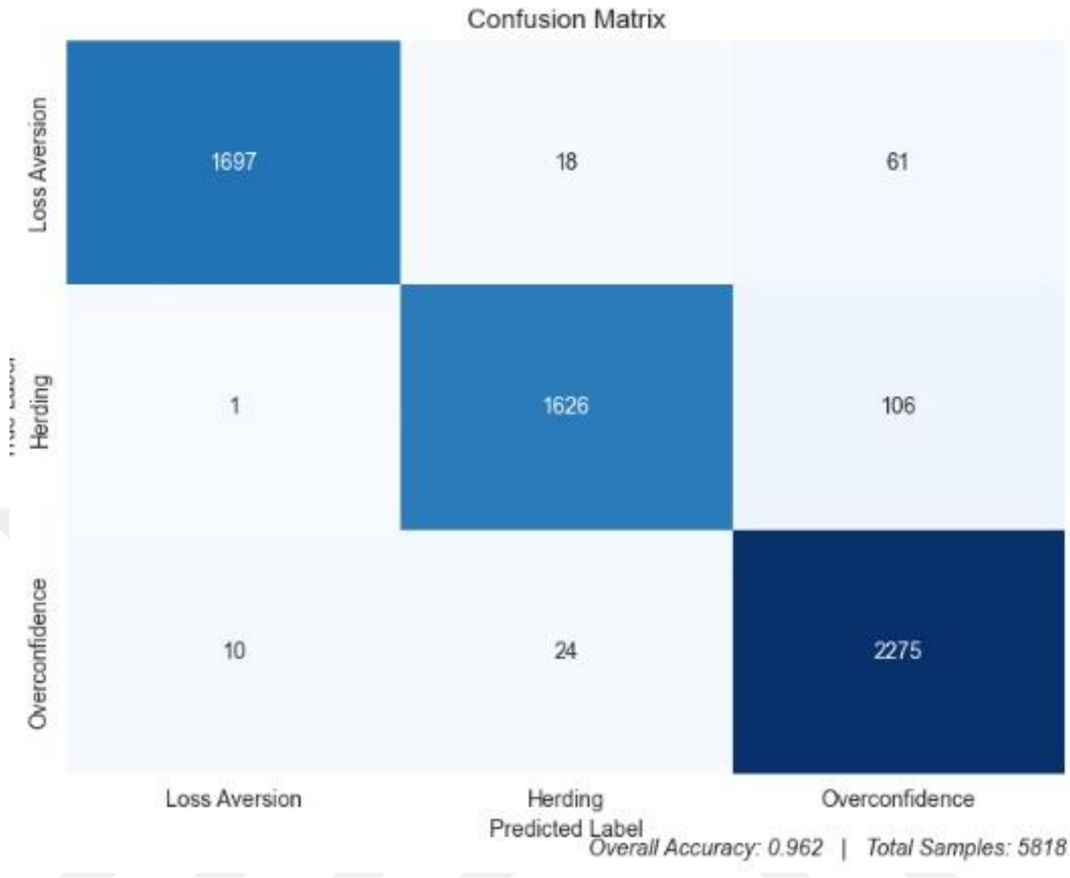
The consistently high ROC-AUC scores of 0.99 across all behavioral bias groups indicate a near-perfect discrimination capability of the model. These exceptionally high scores suggest that the model can almost perfectly distinguish between the different behavioral bias categories, providing compelling evidence for the validity and effectiveness of the behavioral segmentation approach employed in this study. Such high AUC values are a strong indicator of the model's ability to correctly rank positive instances higher than negative ones, showcasing its robust predictive power for identifying investor behavioral biases (Hanley & McNeil, 1982).

4.8. Classification Performance and Error Breakdown

Overall accuracy: 0.962 (5,598 correctly classified out of 5,818).

Confusion matrix (rows = true class, columns = predicted class):

Figure 4.8.1 Confusion Matrix



Class supports: Loss Aversion = 1,776; Herding = 1,733; Overconfidence = 2,309.

Per-class error rates (one-versus-rest):

• Loss Aversion: TP = 1697, FP = 11, TN = 4031, FN = 79 → FPR = 0.0027, FNR = 0.0445.

• Herding: TP = 1626, FP = 42, TN = 4043, FN = 107 → FPR = 0.010, FNR = 0.062.

• Overconfidence: TP = 2275, FP = 167, TN = 3342, FN = 34 → FPR = 0.048, FNR = 0.015.

Macro averages: $\mathrm{FPR} = 0.020, \quad \mathrm{FNR} = 0.040$

Per-class precision/recall/F1:

• Loss Aversion: precision 0.99, recall 0.96, F1 0.97 (support 1,776)

• Herding: precision 0.96, recall 0.94, F1 0.95 (support 1,733)

• Overconfidence: precision 0.93, recall 0.99, F1 0.96 (support 2,309)

• Overall: accuracy 0.96, macro-avg F1 0.96, weighted-avg F1 0.96

This error profile shows a very low false-positive rate for Loss Aversion, a very low false-negative rate for Overconfidence, and the highest miss rate for Herding, which is consistent with the confusion matrix and the class-wise precision and recall reported above.

4.9. Unsupervised Clustering: Summary of Findings

The exploratory k-means and DBSCAN analyses **did not** yield cluster assignments that map cleanly to the three behavioral labels. Solutions lacked stability across parameterizations and did not correspond to the label space. We therefore retained the supervised approach described in Sections 3.1.3–3.1.4 and report performance only for the decision-tree classifier.

4.10. Summary of Key Findings

The comprehensive empirical analysis conducted in this study yields several important findings regarding the classification of behavioral biases among investors:

- **Distinct Behavioral Patterns:** The three behavioral metrics—loss aversion, herding ratio, and average trades per month—successfully capture distinct investment behavior patterns. Each group demonstrates clear separation based on their respective dominant behavioral characteristics. Specifically, the loss aversion group exhibits significantly higher tendencies to hold losing positions (Shefrin & Statman, 1985), the herding group shows a pronounced inclination to follow popular assets (Bikhchandani & Sharma, 2001), and the overconfidence group demonstrates an elevated trading frequency (Barber & Odean, 2001).
- **Statistical Significance of Group Differences:** The statistical significance tests (Kruskal-Wallis H tests) unequivocally confirm that the observed differences between the behavioral bias groups are not due to random variation. This provides robust evidence for the existence of genuinely distinct behavioral segments within the investor population, supporting H2.
- **High Model Performance and Robustness:** The decision tree classification model demonstrated exceptional performance, achieving an overall accuracy of 96% and

ROC-AUC scores of 0.99 for all behavioral bias groups. The consistency of these results across 5-fold cross-validation further supports the robustness and generalizability of the proposed approach, confirming H3.

- **Balanced Distribution of Biases:** The balanced distribution of investors across the three behavioral bias categories (overconfidence being the most prevalent at 39.2%, followed by loss aversion at 30.8%, and herding at 30.0%) suggests that no single bias overwhelmingly dominates the investment population. This highlights the importance of considering multiple behavioral factors when analyzing investment behavior.

These findings collectively contribute significantly to the behavioral finance literature by providing robust empirical evidence for the measurability and predictability of behavioral biases in investment decision-making. Furthermore, the study establishes a reliable and effective framework for behavioral investor segmentation.

4.11. Feature Importance Analysis

The decision tree model also allowed for an analysis of feature importance, revealing which behavioral metrics were most influential in predicting the behavioral biases (Lantz, 2013). The features were ranked as follows:

- **Loss Aversion Score:** Identified as the feature with the highest discriminatory power, indicating its crucial role in differentiating investor groups.
- **Herding Ratio:** Ranked as the second most important feature, signifying its significant contribution to the classification process.
- **Average Monthly Transaction Count:** Placed as the third most important feature, still playing a considerable role in the model's predictions.

Identifying the features that most strongly predict harmful biases enables targeted literacy interventions and platform-level decision aids, which can enhance economic productivity (SDG 8) and inform robust fintech infrastructure (SDG 9).

4.12. Limitations of the Study

Despite the significant contributions of this study, it is important to acknowledge certain limitations that offer avenues for future research.

4.12.1. Data Set Limitations

- **Time Period:** The study utilized data spanning from 2018 to 2022. Consequently, investor behaviors during and after the COVID-19 pandemic, a period marked by unprecedented market volatility and shifts in investor sentiment (Baker et al., 2020), were not included in the scope of this analysis.
- **Geographical Scope:** The FAR-Trans dataset is based on data from a European financial institution. Therefore, the findings of this study may not be directly generalizable to different geographical regions or market structures, as investor behavior can vary significantly across cultures and regulatory environments (Chui & Kwok, 2008).
- **Transaction Details:** The dataset, while rich in transaction records, does not contain information regarding investors' underlying reasons for trades, their specific expectations, or their psychological states at the time of decisions. This absence limits a deeper qualitative understanding of the motivations behind observed behavioral biases (Loewenstein et al., 2001).

4.12.2. Methodological Limitations

- **Causality:** As the study relies on observational data, it is not possible to establish definitive causal relationships between behavioral biases and investment performance. While correlations are identified, further experimental designs would be needed to infer causality (Shadish et al., 2002).
- **Dynamic Analysis:** The study implicitly assumes that investors exhibit a consistent dominant bias over time, overlooking the possibility that investors may display different biases dynamically depending on market conditions or personal circumstances (Baker & Wurgler, 2007).

- **Interaction Effects:** The interactions between different behavioral biases were not examined. In reality, investors may be influenced by multiple biases simultaneously, and their combined effect could be more complex than the sum of their individual parts (Pompian, 2012).

5. DISCUSSION

5.1. Interpretation of Findings

The findings of this study provide strong evidence that individual investors exhibit a heterogeneous structure in terms of behavioral biases, and that these biases can be identified with high accuracy using machine learning methods. The classification performance, achieving a remarkable 96% accuracy rate, demonstrates that behavioral finance theories have tangible manifestations within transactional data. This aligns with a growing body of research utilizing computational methods to understand investor psychology (Shleifer, 2000).

5.1.1. Loss Aversion Findings

The observation that investors in the loss aversion group have an average loss aversion score of 706.36 ± 573.57 indicates that these investors held losing positions for significantly longer periods compared to profitable ones. This finding is in complete alignment with Shefrin and Statman's (1985) disposition effect theory, which posits that investors are reluctant to realize losses and are eager to realize gains. Particularly striking is the median value of the loss aversion metric, which stands at 0.00. This suggests that the vast majority of investors did not realize any losses during the analysis period, strongly supporting Odean's (1998) finding that "investors are unwilling to realize losses." This behavioral tendency can lead to suboptimal investment decisions and prolonged exposure to declining assets, thus hindering wealth accumulation (Glaser & Weber, 2007).

5.1.2. Herding Behavior Findings

Investors categorized within the herding behavior group exhibited an average herding ratio of 0.81 ± 0.30, implying that 81% of their transactions were concentrated in popular assets. This result strongly supports Bikhchandani and Sharma's (2001) theory of informational cascades, which suggests that individuals often disregard their private information and instead follow the actions of others, particularly when information is asymmetrical or ambiguous. The high prevalence of herding behavior observed in this study carries significant implications for market efficiency (Hirshleifer & Teoh, 2003). Such collective behavior can contribute to the formation of price bubbles (Shiller, 2000), as investors may collectively overvalue popular assets, creating an unsustainable surge in prices. Understanding this dynamic is crucial for both individual investors seeking to avoid herd-driven pitfalls and regulators aiming to maintain market stability.

5.1.3. Overconfidence Findings

The overconfidence group constituted the largest segment of investors, accounting for 39.2% of the sample. Furthermore, investors in this group engaged in an average of 2.49 ± 2.95 transactions per month, which directly corroborates the findings of Barber and Odean (2001) regarding the positive relationship between overconfidence and high trading frequency. The elevated trading frequency observed in these investors can be interpreted as a direct manifestation of their excessive belief in their own market forecasting abilities or stock-picking prowess (Daniel et al., 1998). Behavioral finance literature consistently indicates that excessive trading, often fueled by overconfidence, generally leads to lower net returns due to increased transaction costs and poor decision-making (Odean, 1999; Grinblatt & Keloharju, 2009). This highlights a critical area where financial education and behavioral interventions could significantly benefit investors.

6. CONCLUSION

This study successfully demonstrated that behavioral biases exhibited by individual investors in financial markets can be effectively identified and classified using machine learning techniques. Analyses conducted on the FAR-Trans dataset, provided by the University of Glasgow, revealed that 29,090 investors significantly differed in terms of loss aversion, herding behavior, and overconfidence biases.

The primary findings of this research can be summarized as follows: First, investor behavioral profiles showed a heterogeneous distribution, with 30.8% exhibiting loss aversion, 30.0% displaying herding behavior, and 39.2% demonstrating overconfidence. This distribution clearly indicates that investors' decision-making is influenced by diverse psychological factors, consistent with the foundational principles of behavioral finance (Kahneman & Tversky, 1979). Second, the developed decision tree model accurately distinguished these three behavioral groups with an impressive 96% accuracy rate. The model's performance, achieving a 0.99 ROC-AUC value for all classes, unequivocally proves that behavioral biases can be reliably detected from transaction data, showcasing the power of computational methods in this domain (Hastie et al., 2009).

Specific insights into each bias further strengthen these conclusions. Investors in the loss aversion group held losing positions for an average of 706 times longer, strongly supporting Shefrin and Statman’s (1985) disposition effect theory. For the herding behavior group, 81% of their transactions were concentrated in popular assets, aligning with Bikhchandani and Sharma’s (2001) information cascade theory. Lastly, the overconfidence group was the largest segment and exhibited high trading frequency, corroborating the findings of Barber and Odean (2001) regarding the link between overconfidence and excessive trading.

This study makes several significant contributions to the behavioral finance literature. Firstly, by simultaneously examining three core behavioral biases, it demonstrates that investors can be effectively classified based on their dominant biases, moving beyond single-bias analyses. Secondly, it highlights the high performance of the decision tree algorithm in detecting behavioral biases, providing a practical tool for financial analysis and investor profiling. Thirdly, it introduces innovative measurement methods, such as the improved loss aversion metric calculated using the FIFO method, which can be replicated in future research.

Nevertheless, the study acknowledges certain limitations. The dataset's temporal scope (2018-2022) precluded analysis of investor behavior during the COVID-19 pandemic, a period that significantly impacted financial markets and investor psychology (Baker et al., 2020). Furthermore, being an observational study, it could not establish causal relationships between behavioral biases and investment performance. The dynamic nature of investor biases, where individuals might exhibit different biases over time (Baker & Wurgler, 2007), was also beyond the scope of this static analysis.

Future research opportunities are abundant. Longitudinal studies examining the evolution of behavioral biases, comparative analyses of long-term performance across different behavioral groups, and effectiveness evaluations of intervention programs aimed at reducing biases are all promising avenues. Additionally, developing complex models that account for the interaction of multiple biases could provide a more holistic understanding of investor behavior.

In conclusion, this study firmly establishes that behavioral finance theories have concrete manifestations in real transaction data, and that machine learning methods are effective tools for detecting these biases. The insights derived from this research offer valuable guidance for investors to understand their own behavioral tendencies and for financial institutions to provide more effective advisory services. Taken together, the findings position this work as a practical step toward SDG 8 and SDG 9, coupling responsible growth with innovation-ready analytics infrastructure. The systematic detection and analysis of behavioral biases can significantly contribute to a better understanding of financial markets and ultimately enhance investor welfare.

6.1. Recommendations for Future Research

Building upon the insights gained from this study, several promising avenues for future research emerge:

- **Longitudinal Analysis:** Conducting longitudinal studies that examine the evolution of investors' behavioral biases over time would provide valuable insights into their dynamic nature and potential shifts in response to market events or personal experiences (Malmendier & Tate, 2005).
- **Performance Analysis:** A comparative analysis of the long-term investment performance of different behavioral groups could shed light on the financial implications of specific biases and identify which biases are most detrimental to investor returns (e.g., Frazzini & Pedersen, 2014).
- **Multiple Bias Models:** Developing complex models that account for the simultaneous exhibition and interaction of multiple biases in investors' decision-making processes would offer a more nuanced understanding of investor behavior (e.g., Ricciardi & Simon, 2000).
- **Intervention Studies:** Testing the effectiveness of intervention programs designed to mitigate the identified biases could have significant practical implications for financial education and investor protection (e.g., Thaler & Benartzi, 2004).
- **Cross-Cultural Comparisons:** Conducting comparative studies across different cultural contexts would help understand how behavioral biases manifest and vary across diverse investor populations and market structures (e.g., Grinblatt & Keloharju, 2001).

7. REFERENCES

[1] Adil, M., Singh, Y., & Ansari, M. (2021). How financial literacy moderate the association between behaviour biases and investment decision? *Asian Journal of Accounting Research, 7*(1), 17–30. https://doi.org/10.1108/ajar-09-2020-0086

[2] Aktaş, S. (2021). Gelişmekte olan ülkelerde kripto para kullanımının davranışsal finans perspektifinden analizi: Türkiye örneği [Yayımlanmamış yüksek lisans tezi]. Altınbaş Üniversitesi, İstanbul.

[3] Almansour, B. Y., & Arabyat, Y. A. (2017). Investment decision making among Gulf investors: Behavioural finance perspective. *International Journal of Management Studies*, *24*(1). https://doi.org/10.32890/ijms.24.1.2017.10476

[4] Ammann, M., Ising, A., & Kessler, S. (2011). Disposition effect and mutual fund performance. *Applied Financial Economics, 22*(1), 1–19. https://doi.org/10.1080/09603107.2011.595676

[5] Apochi, J., Ahmed, N., Okpanachi, J., & Agbi, S. (2024). Herding bias and financial risk tolerance on individual investment performance in Nigeria: Moderated by financial literacy. *International Journal of Business and Management Review, 12*(3), 1–17. https://doi.org/10.37745/ijbmr.2013/vol12n3117

[6] Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes, 35*(1), 124–140. https://doi.org/10.1016/0749-5978(85)90049-4

[7] Arslan, E., & Boztosun, D. (2022). Demografik faktörler ve davranışsal finans: Bireysel yatırımcılar üzerine bir araştırma. *Turizm Ekonomi ve İşletme Araştırmaları Dergisi*, *4*(1), 30–46.

[8] Asad, H., Khan, A., & Faiz, R. (2018). Behavioral biases across stock market investors: Evidence from Pakistan. *Pakistan Economic and Social Review, 56*(1), 185–209.

[9] Ayaa, M. M., Peprah, W. K., Mensah, M. O., Owusu-Sekyere, A. B., & Daniel, B. (2022). Influence of heuristic techniques and biases in investment decision-making: A conceptual analysis and directions for future research. *International Journal of Academic Research in Business and Social Sciences, 12*(5), 1252–1267. https://doi.org/10.6007/IJARBSS/v12-i5/13339

[10] Bailey, R. E. (2005). *The economics of financial markets*. Cambridge University Press.

[11] Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, *21*(2), 129–151. https://doi.org/10.1257/jep.21.2.129

[12] Baker, S. R., Bloom, N., Davis, S. J., Kost, K. J., Sammon, M. C., & Viratyosin, T. (2020). The unprecedented stock market reaction to COVID-19. *Review of Asset Pricing Studies, 10*(4), 742–758. https://doi.org/10.1093/rapstu/raaa008

[13] Baltussen, G. (2009). Behavioral finance: An introduction. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1488110

[14] Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics, 116*(1), 261–292.

[15] Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. In G. M. Constantinides, M. Harris, & R. M. Stulz (Eds.), *Handbook of the economics of finance* (Vol. 1, pp. 1053–1128). Elsevier. https://doi.org/10.1016/S1574-0102(03)01027-6

[16] Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, 49*(3), 307–343.

[17] Barnett, V., & Lewis, T. (1994). *Outliers in statistical data* (3rd ed.). Wiley.

[18] Barone, R. (2003). *From efficient markets to behavioral finance* (University of Lecce, Economics Working Paper No. 46/24). http://papers.ssrn.com/_id=493545

[19] Becker, G. S. (1976). *The economic approach to human behavior*. University of Chicago Press.

[20] Belsky, G., & Gilovich, T. (2010). *Why smart people make big money mistakes and how to correct them: Lessons from the new science of behavioral economics*. Simon & Schuster.

[21] Ben-David, I., Graham, J. R., & Harvey, C. R. (2013). Managerial miscalibration. *The Quarterly Journal of Economics, 128*(4), 1547–1584. https://doi.org/10.1093/qje/qjt023

[22] Bernard, S., Loos, B., & Weber, M. (2022). The disposition effect in boom and busts markets. *Electronic Journal*.

[23] Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk (L. Sommer, Trans.). *Econometrica, 22*(1), 23–36. (Original work published 1738)

[24] Bikhchandani, S., & Sharma, S. (2000). Herd behavior in financial markets: A review. *IMF Staff Papers, 47*(3), 279–310.

[25] Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth.

[26] Breitmayer, B., Hasso, T., & Pelster, M. (2019). Culture and the disposition effect. *Economics Letters, 184*, 108653. https://doi.org/10.1016/j.econlet.2019.108653

[27] Byrne, A., & Brooks, M. (2008). Behavioral finance: Theories and evidence. *Research Foundation of CFA Institute Literature Reviews, 3*(1), 1–26. https://doi.org/10.2470/rflr.v3.n1.1

[28] Ceylan, A., & Korkmaz, T. (2015). *Sermaye piyasası ve menkul kıymet analizi*. Ekin Kitabevi.

[29] Cheng, Z. (2022). Psychology analysis of investors from the perspective of behavioral finance. In *Proceedings of the 2022 International Conference on Economics, Smart Finance and Contemporary Trade (ESFCT 2022)* (pp. 727–733).

[30] Christie, W. G., & Huang, R. D. (1995). Following the Pied Piper: Do individual returns herd around the market? *Financial Analysts Journal*, *51*(4), 31–37. https://doi.org/10.2469/faj.v51.n4.1918

[31] Chui, A. C. W., & Kwok, C. C. Y. (2008). National culture and life insurance consumption. *Journal of International Business Studies*, *39*(1), 88–101. https://doi.org/10.1057/palgrave.jibs.8400316

[32] Chukwudum, Q. C. (2016). Foundations of expected utility theory and its role in the purchase of insurance. *International Journal of Economics, Commerce and Management*, *4*(6), 42–59. https://ijecm.co.uk/wp-content/uploads/2016/06/463.pdf

[33] Cialdini, R. B. (1998). *Influence: The psychology of persuasion* (Rev. ed.). Quill.

[34] Cici, G. (2012). The prevalence of the disposition effect in mutual funds' trades. *Journal of Financial and Quantitative Analysis*, *47*(4), 795–820. https://doi.org/10.1017/S0022109012000348

[35] Da Silva, S., Da Costa Jr, N., & Mineto, C. (2008). Disposition effect and gender. *Applied Economics Letters*, *15*, 411–416. https://doi.org/10.1080/13504850600706560

[36] Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *The Journal of Finance*, *53*(6), 1839–1885. https://doi.org/10.1111/0022-1082.00077

[37] Daniel, K., Hirshleifer, D., & Teoh, S. H. (2002). Investor psychology in capital markets: Evidence and policy implications. *Journal of Monetary Economics*, *49*, 139–209.

[38] Demirel, E., & Yelkikalan, N. (2020). Karar verme sürecinde yatırımcıları etkileyen bilişsel ve duygusal önyargılar. *Social, Mentality and Researcher Thinkers Journal*, *6*(39), 2583–2593.

[39] Dew, J., & Xiao, J. J. (2011). The financial management behavior scale: Development and validation. *Journal of Financial Counseling and Planning*, *22*(1), 43–59.

[40] Dhar, R., & Zhu, N. (2006). Up close and personal: Investor sophistication and the disposition effect. *Management Science*, *52*(5), 726–740. https://doi.org/10.1287/mnsc.1040.0473

[41] Edgeworth, F. Y. (1881). *Mathematical psychics: An essay on the application of mathematics to the moral sciences*. Kegan Paul.

[42] Elmas, B. (2014). Davranışsal finans. In B. Güngör (Ed.), *Sermaye ve para piyasaları* (pp. 298–322). Atatürk Üniversitesi Açıköğretim Fakültesi Yayını.

[43] Ertuğrul Ayrancı, A. (2020). Bireylerin yatırım kararlarının davranışsal finans yardımıyla incelenmesi: Akademisyenlere yönelik değerlendirme. *Biga İ.İ.B.F. Dergisi, 1*(1), 13–29.

[44] Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* (pp. 226–231).

[45] Fama, E. F. (1965). The behavior of stock-market prices. *Journal of Business, 38*(1), 34–105.

[46] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383–417. https://doi.org/10.2307/2325486

[47] Fama, E. F. (1976). Efficient capital markets: Reply. *The Journal of Finance, 31*(1), 143–145.

[48] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

[49] Feng, L., & Seasholes, M. S. (2005). Do investor sophistication and trading experience eliminate behavioral biases in financial markets? *Review of Finance, 9*(3), 305–351. https://doi.org/10.1007/s10679-005-2262-0

[50] Festinger, L. (1965). *A theory of cognitive dissonance*. Stanford University Press.

[51] Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). SAGE.

[52] Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics, 111*(1), 1–25. https://doi.org/10.1016/j.jfineco.2013.10.005

[53] Cen, L., Hilary, G., & Wei, K. J. (2013). The role of anchoring bias in the equity market: Evidence from analysts' earnings forecasts and stock returns. *Journal of Financial and Quantitative Analysis, 48*(1), 47–76.

[54] Glaser, M., & Weber, M. (2007). Overconfidence and trading volume. *Geneva Risk and Insurance Review, 32*(1), 1–36. https://doi.org/10.1007/s10713-007-0003-3

[55] Grinblatt, M., & Keloharju, M. (2009). Sensation seeking, overconfidence, and trading activity. *Journal of Finance, 64*(2), 549–578.

[56] Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review, 70*(3), 393–408.

[57] Gül, M. (2017). Davranışsal finans açısından altın yatırımcılarının davranışlarının incelenmesi: Gaziantep örneği. *Maliye ve Finans Yazıları, (108)*, 143–165. https://doi.org/10.33203/mfy.313955

[58] Gümüş, F. B., Koç, M., & Agalarova, M. (2013). Bireysel yatırımcıların yatırım kararları üzerinde etkili olan demografik ve psikolojik faktörlerin tespiti üzerine bir çalışma: Türkiye ve Azerbaycan uygulaması. *Kafkas Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, *4*(6), 71–93.

[59] Haigh, M. S., & List, J. A. (2005). A simple test of expected utility theory using professional traders. *Proceedings of the National Academy of Sciences*, *102*(3), 945–948. https://doi.org/10.1073/pnas.0408022101

[60] Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed., Pearson New International Edition). Pearson.

[61] Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.

[62] Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*(1), 29–36. https://doi.org/10.1148/radiology.143.1.7063747

[63] Hanson, J. D., & Kysar, D. A. (1999). Taking behavioralism seriously: Some evidence of market manipulation. *Harvard Law Review*, *112*(7), 1420–1572.

[64] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

[65] He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263–1284.

[66] Hirshleifer, D. (2015). Behavioral finance. *Annual Review of Financial Economics*, *7*, 133–159.

[67] Hirshleifer, D., & Teoh, S. H. (2003). Herd behaviour and cascading in capital markets: A review and synthesis. *European Financial Management*, *9*(1), 25–66. https://doi.org/10.1111/1468-036X.00207

[68] Jain, N., & Kesari, D. B. (2019). Effect of conservatism bias and self-attribution bias on financial risk tolerance of investors. *GAP Gyan—An International Peer-Reviewed Open Access Journal of Social Sciences*, *2*(1), 83–87.

[69] Jaiyeoba, H., Abdullah, M., & Ibrahim, K. (2019). Institutional investors vs retail investors. *International Journal of Bank Marketing*, *38*(3), 671–691. https://doi.org/10.1108/ijbm-07-2019-0242

[70] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. Springer. https://doi.org/10.1007/978-1-4614-7138-7

[71] Joo, B. A., & Durri, K. (2017). Influence of overconfidence, optimism and pessimism on the rationality of the individual investors: An empirical analysis. *Pacific Business Review International, 9*(12), 7–13.

[72] Kaban, L., & Linata, E. (2024). The risk perception as a mediator between herding and overconfidence on investment decision by Gen Z in Indonesia. *Mec-J (Management and Economics Journal), 8*(1), 1–14.

[73] Kahneman, D. (2017). *Hızlı ve yavaş düşün* (O. Ç. Deniztekin, Çev.). Varlık Yayınları.

[74] Kahneman, D., & Riepe, M. W. (1998). Aspects of investor psychology: Beliefs, preferences, and biases investment advisors should know about. *The Journal of Portfolio Management, 24*(4), 52–65. https://doi.org/10.3905/jpm.1998.409643

[75] Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*(3), 430–454. https://doi.org/10.1016/0010-0285(72)90016-3

[76] Kahneman, D., & Tversky, A. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.

[77] Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291. https://doi.org/10.2307/1914185

[78] Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives, 5*(1).

[79] Karakuş, Y. (2022). Davranışsal finans perspektifinden yatırımcı davranışları. *Finansal Araştırmalar ve Çalışmalar Dergisi, 14*(27), 56–72.

[80] Karan, M. B. (2018). *Yatırım analizi ve portföy yönetimi* (5. baskı). Gazi Kitabevi.

[81] Karsh, S. (2018). Investor behavioral finance: Examining its applicability on Egyptian investors. *American Journal of Industrial and Business Management, 8*(11), 2158–2167. https://doi.org/10.4236/ajibm.2018.811143

[82] Katona, G. (1947). *Psychological analysis of economic behavior*. McGraw-Hill.

[83] Katona, G. (1951). *Psychological analysis of economic behavior*. McGraw-Hill.

[84] Katona, G. (1942). *War without inflation: The psychological approach to problems of war economy*. Columbia University Press.

[85] Katona, G. (1947). Contribution of psychological data to economic analysis. *Journal of the American Statistical Association, 42*(239), 449–459. https://doi.org/10.1080/01621459.1947.10501939

[86] Keynes, J. M. (2018). *The general theory of employment, interest, and money*. Palgrave Macmillan.

[87] Kobiyh, M., Amri, A. E., Oulfarsi, S., & Hilmi, Y. (2023). Behavioral finance and the imperative to rethink market efficiency. *Financial Markets, Institutions and Risks*, *7*(4), 38–53. https://doi.org/10.61093/fmir.7(4).38-53.2023

[88] Kocabey, E. (2022, September 1). Bilişsel uyumsuzluk: İnsanlar birbiriyle çelişen inançlara inanmaya nasıl devam edebiliyorlar? *Evrim Ağacı*. https://evrimagaci.org/bilissel-uyumsuzluk-insanlar-birbiriyle-celisen-inanclara-inanmaya-nasil-devam-edebiliyorlar-12268

[89] Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)* (pp. 1137–1145). Morgan Kaufmann.

[90] Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*, 583–621. https://doi.org/10.1080/01621459.1952.10483441

[91] Kukacka, J., & Barunik, J. (2012). Behavioural breaks in the heterogeneous agent model: The impact of herding, overconfidence, and market sentiment. *arXiv preprint* arXiv:1205.3763.

[92] Kumar, S., & Goyal, N. (2016). Evidence on rationality and behavioural biases in investment decision making. *Qualitative Research in Financial Markets*, *8*(4), 270–287. https://doi.org/10.1108/qrfm-05-2016-0016

[93] Lantz, B. (2013). *Machine learning with R*. Packt Publishing.

[94] Lasantha, R., & Kumara, W. (2021). Behavioral biases and the rationality in stock market investments: Evidence from Sri Lanka. *Journal of Business and Technology*, *5*(0), 4. https://doi.org/10.4038/jbt.v5i0.49

[95] Levy, J. S. (1992). An introduction to prospect theory. *Political Psychology*, *13*(2), 171–186.

[96] Liu, Y.-Y., Tseng, V. S., & Tsai, W.-C. (2014). Prospect theory for online financial trading: Large-scale empirical study. *arXiv preprint* arXiv:1402.6393.

[97] Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137.

[98] Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*(2), 267–286. https://doi.org/10.1037/0033-2909.127.2.267

[99] Loppies, L. S., Esomar, M. J. M. J., & Janah, I. N. (2022). Herding behavior, overconfidence, regret aversion bias on investment decisions. *International Journal of*

Economics Social Science Entrepreneurship and Technology (IJESET), 1(5), 345–352. https://doi.org/10.55983/ijeset.v1i5.357

[100] MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 281–297).

[101] Malcıoğlu, G., & Aydın, M. (2016). Borsa İstanbul'da piyasa etkinliğinin analizi: Harvey doğrusallık testi. *Journal of Accounting, Finance and Auditing Studies, 2*(1), 112–123.

[102] Malkiel, B. G. (1973). *A random walk down Wall Street*. W. W. Norton.

[103] Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *The Journal of Finance, 60*(6), 2661–2700. https://doi.org/10.1111/j.1540-6261.2005.00813.x

[104] McDermott, R. (2004). Prospect theory in political science: Gains and losses from the first decade. *Political Psychology, 25*(2), 289–312.

[105] Mill, J. S. (1836). On the definition of political economy and on the method of investigation proper to it. *London and Westminster Review, 26*, 1–29.

[106] Mongin, P. (1997). Expected utility theory. In J. Davis, W. Hands, & U. Mäki (Eds.), *Handbook of economic methodology* (pp. 342–350). Edward Elgar.

[107] Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review, 115*(2), 502–517.

[108] Muermann, A., & Wise, J. (2006). Regret, pride, and the disposition effect. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.930675

[109] Mullainathan, S., & Thaler, R. H. (2000). Behavioral economics. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.245828

[110] Noti, G., Levi, E., Kolumbus, Y., & Daniely, A. (2016). Behavior-based machine-learning: A hybrid approach for predicting human decision making. *arXiv preprint* arXiv:1611.10228.

[111] Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance, 53*(5), 1775–1798.

[112] Odean, T. (1999). Do investors trade too much? *American Economic Review, 89*(5), 1279–1298.

[113] Oran, J. S., & Perek, S. G. (2013). An empirical test of optimism bias in capital budgeting decisions. *Journal of Modern Accounting and Auditing, 9*(2), 287–296.

[114] Pailer, M. K., & Yıldız, B. (2021). *Davranışsal finansın temel psikolojik eğilimleri ve Z kuşağı*. Iksad Publications.

[115] Pikulina, E., Renneboog, L., & Tobler, P. N. (2017). Overconfidence and investment: An experimental approach. *Journal of Corporate Finance, 43*, 175–192. https://doi.org/10.1016/j.jcorpfin.2017.01.002

[116] Piras, L. (2012). How psychology affects decisions in corporate finance: Traditional vs. behavioural approach. *Journal of Governance and Regulation, 1*(4), 76–87. https://doi.org/10.22495/jgr_v1_i4_p6

[117] Pompian, M. M. (2012). *Behavioral finance and investor types: Managing behavior to make better investment decisions*. Wiley.

[118] Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies, 2*(1), 37–63.

[119] Prosad, J. M., Kapoor, S., & Sengupta, J. (2015). Behavioral biases of Indian investors: A survey of Delhi-NCR region. *Qualitative Research in Financial Markets, 7*(3), 230–263. https://doi.org/10.1108/QRFM-04-2014-0012

[120] Qasim, M., Hussain, R. Y., Mehboob, I., & Arshad, M. (2019). Impact of herding behavior and overconfidence bias on investors' decision-making in Pakistan. *Accounting, 5*(2), 81–90. https://doi.org/10.5267/j.ac.2018.7.001

[121] Rakesh, H. (2013). Gambler's fallacy and behavioral finance in the financial markets: A case study of Bombay Stock Exchange. *International Journal of Business and Management Invention, 12*(2), 1–7.

[122] Ricciardi, V., & Simon, H. K. (2000). What is behavioral finance? *Business, Education and Technology Journal*, 1–9.

[123] Robbins, L. (1932). *An essay on the nature and significance of economic science*. Macmillan.

[124] Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20*, 53–65.

[125] Ruben, E., & Dumludağ, D. (2015). İktisat ve psikoloji. In D. Dumludağ, Ö. Gökdemir, L. Neyse, & E. Ruben (Eds.), *İktisatta davranışsal yaklaşımlar* (ss. 31–50). İmge Kitabevi.

[126] Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review, 6*(2), 41–49.

[127] Sanz-Cruzado, J., Droukas, N., & McCreadie, R. (2024). FAR-Trans: An investment dataset for financial asset recommendation. *arXiv preprint* arXiv:2407.08692.

[128] Schneider, F. (2010). The influence of public institutions on the shadow economy: An empirical investigation for OECD countries. *Review of Law & Economics, 6*(3), 441–468.

[129] Schneider, S. (2010, June 29). Homo economicus—Or more like Homer Simpson? *Deutsche Bank Research (Current Issues)*.

[130] Sefil Tansever, S., & Çilingiroğlu, K. H. (2011). Foundations of behavioral finance: Cognitive and emotional biases of decision making. *İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi, 10*(19), 247–268.

[131] Seiler, M., Seiler, V., Traub, S., & Harrison, D. (2008). Regret aversion and false reference points in residential real estate. *Journal of Real Estate Research, 30*(4), 461–474. https://doi.org/10.1080/10835547.2008.12091229

[132] Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

[133] Shahani, R., & Ahmed, S. (2023). Psychological and social factors determining investment decisions in cryptocurrency: Exploring the mediating role of cognitive biases. *Journal of Organisational Studies and Innovation, 9*(4), 25–45. https://doi.org/10.51659/josi.22.159

[134] Shefrin, H. (2002). *Beyond greed and fear: Understanding behavioral finance and the psychology of investing*. Oxford University Press.

[135] Shefrin, H. (2006). *Behavioral corporate finance*. McGraw-Hill.

[136] Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance, 40*(3), 777–790.

[137] Shiller, R. J. (2000). *Irrational exuberance*. Princeton University Press.

[138] Shleifer, A. (2000). *Inefficient markets: An introduction to behavioral finance*. Oxford University Press.

[139] Simon, H. A. (1947). *Administrative behavior: A study of decision-making processes in administrative organization*. Macmillan.

[140] Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69*(1), 99–118. https://doi.org/10.2307/1884852

[141] Simon, H. (1955). A behavioral model of choice. *The Quarterly Journal of Economics, 69*(1), 99–118.

[142] Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management, 45*(4), 427–447.

[143] Statman, M. (1999). Behavioral finance: Past battles and future engagements. *Financial Analysts Journal, 55*(6), 18–27.

[144] Taner, T., & Kayalıdere, K. (2002). 1995–2000 döneminde İMKB'de anomali araştırması. *Yönetim ve Ekonomi Dergisi, 9*(1), 1–24.

[145] Tekin, B. (2016). Behavioral biases as an effective factor for the firm financial decision-making: A literature review. *Journal of Social Science Research, 10*(2), 2103–2115. https://doi.org/10.24297/jssr.v10i2.4748

[146] Tekin, B. (2018). Overconfidence and effect on firm financial decisions in the context of behavioral corporate finance. *ELK Asia Pacific Journal of Finance and Risk Management, 9*(4). https://doi.org/10.31511/eapjfrm.2018.v09i04004

[147] Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science, 4*(3), 199–214. https://doi.org/10.1287/mksc.4.3.199

[148] Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making, 12*(3), 183–206. https://doi.org/10.1002/(SICI)1099-0771(199909)12:3<183::AID-BDM318>3.0.CO;2-F

[149] Thaler, R. H. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization, 1*(1), 39–60.

[150] Thaler, R. H. (2015). *Misbehaving: The making of behavioral economics*. W. W. Norton & Company.

[151] Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy, 112*(S1), S164–S187. https://doi.org/10.1086/380085

[152] Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

[153] Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.

[154] Virigineni, M., & Rao, B. M. (2017). Contemporary developments in behavioral finance. *International Journal of Economics and Financial Issues, 7*(1), 448–459.

[155] Von Neumann, J., & Morgenstern, O. (2021). *Theory of games and economic behavior* (60th anniversary commemorative ed.). Princeton University Press. (Original work published 1944)

[156] Weber, M., & Camerer, C. F. (1998). The disposition effect in securities trading: An experimental analysis. *Journal of Economic Behavior & Organization, 33*(2), 167–184.

[157] Wibowo, M., Indrawati, N., & Aisjah, S. (2023). The impact of overconfidence and herding bias on stock investment decisions mediated by risk perception.

International Journal of Research in Business and Social Science, 12(5), 174–184. https://doi.org/10.20525/ijrbs.v12i5.2663

[158] Wijayanti, D., Suganda, T. R., & Thewelis, F. S. (2019). Gambler's fallacy as behavioural bias of young investor. *Journal of Business and Behavioural Entrepreneurship, 3*(2), 72–80.

[159] Xiao, J. J. (2016). *Handbook of consumer finance research*. Springer.

[160] Yang, W. (2016). Survey on behavioral finance theory. In *Proceedings of the 2016 International Conference on Management Science and Innovative Education* (pp. 126–129). https://doi.org/10.2991/msie-16.2016.29

[161] Yi, Z., Wang, Y., & Chen, Y. (2021). Financing an agricultural supply chain with a capital-constrained smallholder farmer in developing economies. *Production and Operations Management, 30*(7), 2102–2121. https://doi.org/10.1111/poms.13357

[162] Zahera, S., & Bansal, R. (2019). A study of prominence for disposition effect: A systematic review. *Qualitative Research in Financial Markets, 11*(1), 2–21. https://doi.org/10.1108/qrfm-07-2018-0081

[163] Zeelenberg, M., van den Bos, K., van Dijk, E., & Pieters, R. (2002). The inaction effect in the psychology of regret. *Journal of Personality and Social Psychology, 82*(3), 314–327. https://doi.org/10.1037/0022-3514.82.3.314

[164] Zhang, Y., & Zheng, X. (2015). A study of the investment behavior based on behavioral finance. *European Journal of Business and Economics, 10*(1). https://doi.org/10.12955/ejbe.v10i1.557

[165] Şamandar, Ö., & Çömlekçi, İ. (2019). Kişilik tiplerinin davranışsal finans kapsamında değerlendirilmesi. *Trakya Üniversitesi Sosyal Bilimler Dergisi, 21*(2), 557–575.

[166] Şentürk, F., & Fındık, H. (2014). Rasyonel karar alan ekonomik birimin risk altında verdiği kararlara davranışsal yaklaşım: Kahneman–Tversky beklenti teorisi perspektifinden eleştirel bir bakış. *Öneri Dergisi, 11*(42), 127–139. https://doi.org/10.14783/od.v11i42.5000065510